


```

Db 79 SAEPPESNKKDLTVFQSVKHEQKLDGCGGYMKLLSSSTDQKKFGGDTPTYSIMFGPDIC 138
QY 121 GPCTKKVHVIFNKGKLVNKLKIDCKDETHLYTLVIRPDNTYEVKIDNSOVESGSL 180
Db 139 GYSTKKVHALNVDNTNHLKKEVPCETDQTHVYLVIRPDATYSILINDVSKQTSGLY 198
QY 181 DMDWFLPPKKIKDPAKDPEDWDERAKIDDPDTSKPDWDK - PEHPTDPDAKPPEDWDEE 239
Db 199 TDWDLPPKKIKDPAKDPEDWDEKEYIPDPEKKPEGYDDIPKEIPDPAKPPEDWDE 258
QY 240 MDGEWPPVTONPEYKGEWKPRQIDNPYKGTWHPIDNPESPDPSIYAYDNFGVLGL 299
Db 259 EDGEWAPTIANPEYKGPWKPKIKNPYKGWKAPMIDNPDKDDPEIYVYNLKYVGI 318
QY 300 DLQVKSGLTFIDNDEAYAEFGNETGWVTKAAEKQMKDKQDEORLKEEEDKKR 359
Db 319 ELQVKSGLTFIDNDEAYAEFGNETGWVTKAAEKQMKDKQDEORLKEEEDKKR 375
QY 360 KEDEAEKDEKDEDEDEDEDEKE---EDEDVPGQAKDEL 400
Db 376 PADSDADEDDDDADTEGEDGESKSAEDSAEDV---HDEL 415

RESULT 2
US-08-296-362-2
; Sequence 2, Application US/08296362
; Patent No. 5691306
; GENERAL INFORMATION:
; APPLICANT: Bergeron, John J.M.
; APPLICANT: Thomas, David Y.
; APPLICANT: Wada, Ikuo
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; TITLE OF INVENTION: PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
; TITLE OF INVENTION: PROTEIN PRODUCTION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,362
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deehr, Manya S.
; REGISTRATION NUMBER: 37,120
; REFERENCE/DOCKET NUMBER: 690066.401C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-296-362-2

```

Query Match 36.2%; Score 799; DB 1; Length 593;
 Best Local Similarity 34.9%; Pred. No. 1.4e-63;
 Matches 182; Conservative 72; Mismatches 133; Indels 134; Gaps 16;

```

QY 4 VYKQEQFLDGGWTSRWIESKHKSD-----FGKVLSSGKFGYGDDEK-----DKGLQTS 52
Db 70 VYFADSF--DRGTLSGWLISKAKKDDTDEIAKY---DGKWEVDMEKTKLPKGKGLVLM 124
QY 53 QDARFYALSASF-EPPSNKQGTLLVQVTVHQENIDCGGGYVVKLFPPN-----SLDQTMHG 107
Db 125 SRAKHHAISAKLNKPPFLDKPLIVQYEVANFQNGIECGGAYVVKLLSKTPELNLDQ--FHD 182
QY 108 DSEYNINFGPDICGPGTKVHVIFNKGKLVNKLKIDCK-----DDEFTHLTYL 157
Db 183 KTYTTFMGDPDKCE-DYKLFHFRHNKPKTGYEEKHAKRPDADLTYTDTKKTHLYTL 241
QY 158 IVRPDNTYEVKIDNSOVESGSLDDWDFLPP-----KKIKDPDASKPDWDERAKIDDPD 213
Db 242 ILNPDNSFEILVQDSIVNSGNLLN--DMTPPVNPSREIEDPEDQKPEDWDERPKIPDPA 299
QY 214 SKPEDW-----DKPEHPTDPDAKPPEDWDEMDGEWEP-----246
Db 300 VKPDDMNEDAPAKIPDEEATKPDGWLDDPEYVPPDPAEKPDWDEMDGEWEPQIAMP 359
QY 247 -----PVIONPEYKGEWKPRQIDNPYKGTWHPIDNPESPDPSIYAYDN 293
Db 360 KCESAPCGGVQWRPMIDNPYKGWKPKPMIDNPYQIWKPKIPNPDFTEDLEPFWKTP 419
QY 294 FGVGLDLQVKSGLTFIDNDEAYAEFGNETGWVTKAAE-----KOMKDKQD 345
Db 420 FSAIGLELWSMTSDIFFDNFIVCGDRRVVDWANDGWLKKAADGAAEPGVVGQMIKAAE 479
QY 346 E-----EORLKEEEDK-----KR 359
Db 480 ERPLMWVYVLTVALPVFLVIFSCCGKQSSPVEYKKTDAPODPVKEEKEEKEEDKDG 539
QY 360 KEDEAEKDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
Db 540 DEEESEKLEEKQKSDAEEDGGTASQEDDRPKAEDEI 580

RESULT 3
US-08-675-816-6
; Sequence 6, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND US
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-816-6

Query Match      29.1%; Score 641.5; DB 4; Length 542;
Best Local Similarity 31.9%; Pred. No. 1.7e-49;
Matches 162; Conservative 55; Mismatches 162; Indels 129; Gaps 17;

QY 6 FKQFLDGDGWTGSRWIESKHSKDFGKVLSSGKPYGDEKDKGLQTSQDARFVALSASF- 64
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 35 FYSEF--EDSFEGRWTLISA-KDDY-KGEWKHAKSEGHDE--YGLLVSEKARKYAVIKELD 88
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 EPESNKQTLVQFTVKEHQNIQCGGYVK-LFPNSLDQI--DMHGSEYNIMFGPDICG 121
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 89 EPATLKDGTIVLQFTRFQNGLECGGAYLKLRLPQEAQWPKDNDSPYSIMFGPKCG 148
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 PGTKKHVHVIENYKG--KNVLINKDIR----CKDDEFHLYTLIVRPDNTVEVKIDNSQVE 175
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 149 -ATNKVHFILKHKPKSGEYIEHHLKYPSPVPSDKLTHVYTAILKPDNELRLVDGEEK 207
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 176 SGSLEDDWDFLPP-----KTKDPDASKPEDWDERAKIDDPDTSKPEDW----- 219
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 208 KANFLSSDDEFPPLVPAKTIPTDPPDKKPEDWDERAKIPDNVAVKPDWDEADAPMEIVDED 267
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 220 -----DKPEHIDPDPAKPEDWDEEMDGEWEPVQIONPE----- 253
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 268 AEKPEGWLDEPEIDPDPAKPEDWDEEDGEWEPVQIONPE----- 327
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 254 --YKGEWKPRQIONPYKGTWIIPEIDNPEY----SPDPSIYAYDNFVGLGLDLWQVKG 307
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 328 PAYKGKHAPLIDNPYKGLWIKQEIIPNPNYFELEKPD-----FEPTAAVGEIWTWQDG 382
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 308 TIFDNFLITNDEAYAEFGNETGWTKAAEKQMKDKOD----- 345
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 ILRGHILIADEKVAESLROTANKPKPDAEKEKQAEADAAGSGLAGFKVFDLLYQV 442
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 346 EQRLKEEEDK----- 368
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 443 ADIPFSEHKDKLIDIEKEKOPNLITIGILSVIVVFTVLFKILFGGKKPAKVEEKPA 502
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 369 EDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 396
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 PAAETSKKESSGEKAENEKEDAAAAA 530
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
US-08-990-114-3
; Sequence 3, Application US/08990114
; Patent No. 5932475
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,114
; FILING DATE: Herewith
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0451 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 714 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 128842
; US-08-990-114-3

Query Match      8.4%; Score 184.5; DB 2; Length 714;
Best Local Similarity 23.9%; Pred. No. 2.5e-08;
Matches 59; Conservative 26; Mismatches 39; Indels 123; Gaps 7;

QY 187 PPKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPDPDAKPKPEDWDEMDGEWEP 246
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 PAKGAKGNKAKKEDSDEDEDDDDDDSDSD-----EEDDED-BEEP 171
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 247 PVTONPEYKGEWKPRQIDNPDKYKGTWIIPEIDNPEYSPDPSIYAYDNFVGLDLWQVKS 306
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 172 PV-----KG-----KQ----- 178
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 307 GTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKDKODEORLKEEEDKKKE----- 361
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 179 GKVA-----AAPASEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 218
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 362 -----EEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 393
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 219 TPAKGKKAPAKVVPVAKNVAEEDDDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 278
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 394 GQAKDEL 400
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 279 GKRKKEM 285
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-08-257-073-5
; Sequence 5, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994

```

COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Review #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-728-323A-2

Query Match 8.0%; Score 175.5; DB 2; Length 1162;
Best Local Similarity 24.4%; Pred. No. 3.le-07;
Matches 62; Conservative 45; Mismatches 88; Indels 59; Gaps 12;

Qy	186	LPFKKIKDDPASPEDOWERAKIDDT	---DSK-----PEDWDKPEHI-----	PDPAK	231
		: :	: :		
Db	134	LPP--TQSPSS-----QRPLSSPTGRPDSTPMRPP	PSOQTTPPHSPHTTTPPP	PEPSK	185
		: :	: :		
Qy	232	KPED-----WDEMDGWEPP-----PVQNPEV-----	KGSMKPRQIDNPY	268	
		: :	: :		
Db	186	SSPDSLAPSTLRSLRRLSSPOGPGSTLNPICGSPVSPRCD	FANRSVYPWATSPY	245	
		: :	: :		
Qy	269	KGWTHPEIDNPYSPP-----DPSIYAY-DNFGVLGLDLMQVKS	GTITFDNFLTIN	317	
		: :	: :		
Db	246	VGS--SSDGDTPRQPPTSPISIGSSSPSEGSWGDDTAMLVLLAE	ATAEASAKNEKESN	303	
		: :	: :		
Qy	318	DEAVAEFGNETGVTKAEQKMKQDEORLKEEEDK	KRKEEAEADKEDDEDKED	377	
		: :	: :		
Db	304	NQA--GEGNGNETSKESQVDKDDNDKNKDEE	EQETDEDEDEDEDEDEDEDE	362	
		: :	: :		
Qy	378	EDEEDKEEDEED	391		
		: :	: :		
Db	363	EDDEEDDEED	376		
		: :	: :		

```

RESULT 7
US-08-574-959A-9
; Sequence 9, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jung, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```



```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
;
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 905 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-574-959A-9

```

```

Query Match      7.4%; Score 163; DB 2; Length 905;
Best Local Similarity 26.3%; Pred. No. 2.9e-06;
Matches 66; Conservative 36; Mismatches 85; Indels 64; Gaps 10;

QY 158 IVRPNTYEVKIDNSQVSGSLEDWD-----FLPPKKIKDPDASKPDPWDERAKID 209
      : || : : || || || || || || || || || || || || || || || ||
Db 526 VPRPAFVHYDKEEASDVEI-SLESDDSDVVIVPEGLPPLPPPPSGATP---PPIAPTG 581
      : || : : || || || || || || || || || || || || || || || ||
QY 210 DPTDSKQ-EDWDKPEHLPDPDAKKPEDWDEMDGEWPPVYIQNYEYKGEWKPRQIDNPDY 268
      : || : : || || || || || || || || || || || || || || || ||
Db 582 PPTASPVPVPAKEEPEELPAAGPLPP-----PPP----- 611
      : || : : || || || || || || || || || || || || || || || ||
QY 269 KGTWIHPIDINPEYSPDPSTIAYDNFG-----VLGLDLQWVKSCTIFDNFLITNDAAVAE 323
      : || : : || || || || || || || || || || || || || || || ||
Db 612 -----PPVPVGVYLPVPPQLVPECTPGGGGPPALEEDL-----TVIN--INSSDEEEE 658
      : || : : || || || || || || || || || || || || || || || ||
QY 324 EFGNETGVTKAAFKOMKQDEQRLEKEEEDKKRKEEAEADKDEDEKDEE---D 380
      : || : : || || || || || || || || || || || || || || || ||
Db 659 EGDEEEEEEEEEEEEEEEEEEDFEEDEDEEYFEEEEEEEEEEEEEEEEEGELE 718
      : || : : || || || || || || || || || || || || || || || ||
QY 381 EEDKEEDEED 391
      : || || || || || || || || || || || || || || || ||
Db 719 EEEEEDEEEE 729
      : || || || || || || || || || || || || || || || ||

```

```

RESULT      8
US-08-574-959A-7
; Sequence 7, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207

```

```

.. REFERENCE/DOCKET NUMBER: DFN-0008
..
.. TELECOMMUNICATION INFORMATION:
..
.. TELEPHONE: (617)227-7400
..
.. TELEFAX: (617)227-5941
..
.. INFORMATION FOR SEQ ID NO: 7:
..
.. SEQUENCE CHARACTERISTICS:
..
.. LENGTH: 1135 amino acids
..
.. TYPE: amino acid
..
.. TOPOLOGY: linear
..
.. MOLECULE TYPE: protein
US-08-574-959A-7

```

Query Match	7.4%;	Score 163;	DB 2;	Length 1135;
Best Local Similarity	26.3%;	Pred. NO. 4e-06;		
Matches 66;	Conservative 36;	Mismatches 85;	Indels 64;	Gaps
Qy 158	IVRPDNTYEVKIDNSQVESGSLEDDWD-----FLPPKKIKDPDASKPDWDERAKID 209			
Db 756	VPRAEAVHYDKEEASOVEI-SLESDDSVVIVPEGLPPLPPPPSCATP---PPIAPTG 811			
Qy 210	DPTDSKP-EDWDKPEHLPDPDAKKPDWDEMDGWEVPPVIQNPEYKGEWKPRQINDPY 268			
Db 812	PPTASPPVPAKEEPEELPAAGPLPP-----PPP-----841			
Qy 269	KGTHIPEINDPEYSPDPSIAYDNEFG-----VLGLDWQVKSCTIFDNLITNDYAE 323			
Db 842	-----PPPVFGPVXLPPLVQPVPECTPGGGGPPALEEDL-----TWIN--INSSDEEEE 888			
Qy 324	EFGNETGWTVKAAFKOMKDKODEQRLLKEEEDKKKEEAEADKEDDEKDKODEE---D 380			
Db 889	ECDEEEEEEEEEEEEEEEEDFEEDEDEEYFEEEEEEEFEEFEESGELE 948			
Qy 381	EEDKEDEEED 391			
Db 949	EEEEDEEEEE 959			

RESULT 9
 US-09-022-983-5
 Sequence 5, Application US/09022983
 Patent No. 6159731
 GENERAL INFORMATION:
 APPLICANT: Yang, Xiaolu
 APPLICANT: Khosravi-Far, Roya
 APPLICANT: Chang, Howard Y.
 APPLICANT: Baltimore, David
 TITLE OF INVENTION: DAXX, A NOVEL FAS-BINDING
 TITLE OF INVENTION: PROTEIN THAT ACTIVATES JNK AND APOPTOSIS
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: Fastseq for Windows
 CURRENT APPLICATION DATA:
 - APPLICATION NUMBER: US/09/022,983
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/037,919
 FILING DATE: 12-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/051,753
 FILING DATE: 26-JUN-1997
 ATTORNEY/AGENT INFORMATION:

```

; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: M0656/7036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 740 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-09-022-983-5

```

```

Query Match 6.8%; Score 151; DB 4; Length 740;
Best Local Similarity 28.4%; Pred. No. 2.6e-05;
Matches 44; Conservative 23; Mismatches 40; Indels 48; Gaps 5;

QY 248 VIONPEYKGEWKPRQIDNPDKYKGTWHPIDNPESPD-----PSIYAYDNFGVLGLDLWQ 303
Db 381 MLODKSEGEKRRK---ARLQCTSH-SADTPEASLDSCEGPS----- 420

QY 304 VKSGTIFDNLITNDAYAEFGNETGWVTKAAEKQKQDEEQLKKEEEDKKRKEE 363
Db 421 -----GMSQGCPSASRAETDDDEES--DEEEEEEEEE 456

QY 364 EARDKEDDEKDEDEDEKDEDEEDVPGQAKD 398
Db 457 EAYDSEEDLEQMQBQEDDEDEDEEBAAGKD 491

```

RESULT 10

```

US-08-935-855-20
; Sequence 20, Application US/08935855
; Patent No. 6066485
; GENERAL INFORMATION:
; APPLICANT: Guthridge, Mark
; APPLICANT: Basilio, Claudio
; TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
; TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Floor
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,855
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-935-855-20

Query Match 6.8%; Score 150.5; DB 3; Length 546;
Best Local Similarity 24.4%; Pred. No. 1.9e-05;
Matches 71; Conservative 31; Mismatches 94; Indels 95; Gaps 12;

QY 158 IVRPDNTYEVKIDNSQVESGLDWDWFLPPKKIKD--PDASKP-EDWDERAKIDDPDTS 214
Db 78 IIKQKAYKQKLEQKALEDAFLAIDAKLTTEEVIKELAQIAGRPTEDEDEKXVADEDDV 137

QY 215 KPED-----WDKPEHFPDPAKKPEDWDEE-----MDGEWEPPIQN 251
Db 138 DNEEAALLHEEATMTIEELLTRYGQCHKGPHSKSGGTGEEPSQGLNGE-----AG 191

QY 252 PEYKGEWKPRQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFGVLGLDLWQVKSQTIFD 311
Db 192 PEDSTRETSQENGPIAK-----AYTGES----- 215

QY 312 NFLITNDEAYAE-----EFGNETWGVTK-----AAEKQKQDEEQLKKEE 354
Db 216 -----SNSRGTEAGVQGEPIPTGEAGPSCSSASDKLPRVAKSKFTFDESEDEAE 271

QY 355 EDKKRKEEE---AEDKEDDEKDEDEDEDEEEDKEEEDDEE-DVPG-QAKDE 399
Db 272 EDSEECSEEDGYSSAEANEDEDDTEAEEDDEEEEMMYPGMEGRKE 322

```

RESULT 11

```

US-08-741-134-6
; Sequence 6, Application US/08741134
; Patent No. 5861498
; GENERAL INFORMATION:
; APPLICANT: Litwack, Gerald
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandes-Alnemri, Teresa
; TITLE OF INVENTION: IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING
; TITLE OF INVENTION: METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5861498ris
; STREET: One Liberty Place - 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 3.11
; SOFTWARE: WordPerfect for Windows 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,134
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/007,163
; FILING DATE: 01-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-2090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:

```

```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 411 amino acids
;
; TYPE: amino acid
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-08-741-134-6

```

Query Match	6.8%;	Score 149;	DB 2;	Length 411;
Best Local Similarity	21.2%;	Pred. No. 1.7e-05;		
Matches	49;	Conservative 52;	Mismatches 82;	Indels 48; Gaps
QY	170	DN SQVSGSLDDOWLEPPKKIKRDPASKPDEWDERAKIDDPDTSKPEDWKPEHP-DP 228		
DB	72	DEDEIDESSEEEEEKTQKKK-----SKGKKAESSEDEDEDDDEDFQESVLTLS 127		
QY	229	DARKPEDWDEMDGEPVPIQNPYKGEWKPKRQIINDPYKGTWIIHPEIDNPSPDPST 288		
DB	128	EAQYQOSLDLTITPEEVQPIVTGSA-----ISLSGNVYKHFPDTP----- 169		
QY	289	YAYDNFVGLGLDLQWKSGTIFDNELITNDEAVAEFGNETGVTAKAAQKMKDDEQ 348		
DB	170	-----MGVEGD-----EDEADIYDSEDYDLTPDEDEIIGDDMD-- 205		
QY	349	RLKEEEDKKRKEEAEDKEDDEDDEDEEEDKEEEDVPGQAKDE 399		
DB	206	-LDDEEEVRIEVEDEEDNDGEEOEDEEEEOBKREVKEPKKSKE 255		

```

RESULT 12
US-09-214-564A-2
; Sequence 2, Application US/09214564A
; Patent No. 6150515
; GENERAL INFORMATION:
; APPLICANT: Sharp, Phillip A.
; APPLICANT: Zhou, Qiang
; TITLE OF INVENTION: TAT-SF: Cofactor For Stimulation Of Transcriptional
; FILE OF INVENTION: Elongation By HIV-1 TAT
; FILE REFERENCE: M0656/7042
; CURRENT APPLICATION NUMBER: US/09/214,564A
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/021,218
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/033,152
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: PCT/US97/11713
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-564A-2

```

```

Query Match          6.6%  Score 146;  DB 4;  Length 754;
Best Local Similarity 25.0%;  Pred. No. 7.5e-05;
Matches 59;  Conservative 45;  Mismatches 88;  Indels 44;  Gaps 9;

QY 169  IDNSQVSGSLEDDWDL-----PPKIKDPDASKPDEDRAKIDDPDTSKPEDW 219
      :  |  |  |  :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 507  LKDCENGLAKESDDLKKESEEVGPTKUSEDDSEKSEDECSQSEBDGSGRFEFE 566
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 220  DKPEHIDPAKPDDEMDGMEPPVIONPEYKGEWKPRQIDNPDKYKGTWTHPEIDN 279
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 567  NGLKDLDEGSEKELHENVLDKLEENDSENSEFDDGSEKVLDE---EGSEREFEDS 623
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 280  PEYSPORSIYADNFGVLGLDLWQVSGTIFDNFL-ITNDAYAEFEFGNETWGTAKAEK 338
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 624  DEKEEEEDTb-----KVFDDSEKDEKBEYADEKLE-----AADK 660
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 339  QMKDKQDEQRKKEEEDKKRKEEAEDKEDDEDKDE--DEEDEEKEDEEDED 391

```

[illegible]

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 6013-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-114-555A-8

Query Match	6.5%	Score 143.5;	DB 2;	Length 240;
Best Local Similarity	47.5%	Pred. NO. 2.5e-05;		
Matches	28;	Conservative	16;	Mismatches 14;
				Indels 1;
				Gaps 1;

QY 34.3 QDEQRKKEEDKKRKEEERAEADKDEDKDDEDEEDEKE-EDDEEDVPGAKDEL 400
| : ||||:::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 201 KMIPEVFTTEEEEEEEEEEEEEEDEDEEKKEDEEEEEKDAEKEEEEAPEGEKEDL 159

RESULT 14
US-08-559-397A-14
Sequence 14, Application US/08559397A
Patent No. 6083713
GENERAL INFORMATION:
APPLICANT: Manly, Susan P.
APPLICANT: Kozlowski, Michael R.
APPLICANT: Neve, Rachael L.
TITLE OF INVENTION: CLONING AND EXPRESSION OF
TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:-
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:02:05 ; Search time 52.92 seconds

(without alignments)
575.771 Million cell updates/sec

Title: US-09-807-148-3

Perfect score: 2206

Sequence: 1 EPVVFKEQFLDGDGWTSRW.....EEDKEDEEDVPGQAKDEL 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2206	100.0	417	1 A37047	calreticulin precu
2	2099.5	95.2	416	1 S06763	calreticulin precu
3	2091.5	94.8	416	2 JH0819	calreticulin precu
4	2070.5	93.9	418	1 A34154	calreticulin precu
5	2059	93.3	400	2 S43376	calreticulin, brai
6	1837.5	83.3	411	2 S29129	calreticulin precu
7	1775.5	80.5	419	2 S71343	calreticulin precu
8	1768.5	80.2	384	2 S29130	calreticulin (clon
9	1687	76.5	405	1 JH0795	calreticulin precu
10	1641.5	74.4	421	2 S36799	calreticulin precu
11	1461.5	66.3	406	2 A56637	calreticulin homol
12	1400.5	63.5	395	2 S25851	calreticulin precu
13	1245	56.4	336	2 A32507	41K larval antigen
14	1206.5	54.7	415	2 T10172	calreticulin - cas
15	1201	54.4	421	2 S58170	calreticulin precu
16	1175	53.3	412	2 T05703	calreticulin - bar
17	1175	53.3	415	2 T05705	calreticulin - bar
18	1171.5	53.1	416	2 T16968	calreticulin call
19	1165.5	52.8	389	2 T03691	calreticulin - com
20	1161.5	52.7	416	2 T14354	calreticulin - bee
21	1150.5	52.2	425	2 C96605	calreticulin (Ortl
22	1142.5	51.8	444	2 H86324	hypothetical prote
23	1086.5	49.3	393	1 A48573	calreticulin autoa
24	992	45.0	422	2 T07841	probable calreticu
25	810	36.7	591	2 B54354	calnexin precursor
26	803	36.4	591	2 C54354	calnexin precursor
27	801	36.3	592	2 I53260	calnexin - human
28	799	36.2	592	2 A46673	calnexin precursor
29	799	36.2	593	1 A37273	calnexin precursor

30	796	36.1	622	2 S71342	calnexin precursor
31	718	32.5	611	2 A53418	calnexin precursor
32	701	31.8	611	2 A54086	calnexin-t - mouse
33	676	30.6	546	2 T06415	calnexin - soybean
34	673.5	30.5	582	2 A46637	calnexin homolog S
35	658	29.8	530	2 JN0597	calnexin-like prot
36	628.5	28.5	560	2 S56142	calcium-binding pr
37	618.5	28.0	619	2 S40938	hypothetical prote
38	606	27.5	532	2 T49873	calnexin homolog -
39	590	26.7	540	2 T10892	probable calnexin
40	587	26.6	356	2 A46164	calnexin - human (
41	551	25.0	428	2 T03251	calnexin - maize (
42	500.5	22.7	297	2 S70552	calnexin homolog C
43	385.5	17.5	43	2 S29347	calnexin homolog Y
44	219	9.9	792	2 T42963	hypothetical prote
45	193	8.7	798	2 T33022	hypothetical prote

ALIGNMENTS

RESULT: 1

A37047

calreticulin precursor - human

N:Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding pr
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Feb-2000
C:Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075
R:McCaulliffe, D.P.; Yang, Y.S.; Wilson, J.; Sontheimer, R.D.; Capra, J.D.
J. Biol. Chem. 267, 2557-2562, 1992

A:Title: The 5'-flanking region of the human calreticulin gene shares homology with t
A:Reference number: A42330; MUID:92129342
A:Accession: A42330

A:Molecule type: DNA

A:Residues: 1-417 <MC2>

A:Note: sequence extracted from NCBI backbone (NCBIN:78537; NCBI:78536)
R:McCaulliffe, D.P.; Lux, F.A.; Lieu, T.S.; Sanz, I.; Hanke, J.; Newkirk, M.M.; Bachin

J. Clin. Invest. 85, 1379-1391, 1990

A:Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/

A:Reference number: A37047; MUID:90237213

A:Accession: A37047

A:Molecule type: mRNA

A:Residues: 1-417 <MCC>

A:Cross-references: GB:M32294; NID:9337486; PIDN:AAA36582.1; PID:9337487
R:Roach, L.A.; Haselby, J.A.; Meilof, J.F.; Smeenk, R.J.; Unnasch, T.R.; Greene, B.

J. Immunol. 147, 3031-3039, 1991

A:Title: Characterization of the autoantigen calreticulin.

A:Reference number: A46452; MUID:92013129

A:Accession: A46452

A:Molecule type: mRNA

A:Residues: 1-417 <ROK>

A:Cross-references: GB:M84739; NID:9179881; PIDN:AAA51916.1; PID:9179882
A:Note: sequence extracted from NCBI backbone (NCBIN:60749; NCBI:60750)

R:Lies, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.

J. Clin. Invest. 82, 96-101, 1988

A:Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence

A:Reference number: A28812; MUID:88273610

A:Accession: A28812

A:Molecule type: protein

A:Residues: 18-41 <LIE>

A:Note: 18-Ala was also found

R:Dupuis, M.; Schaefer, E.; Krause, K.H.; Tschopp, J.

J. Exp. Med. 177, 1-7, 1993

A:Title: The calcium-binding protein calreticulin is a major constituent of lytic gra

A:Reference number: PH1525; MUID:93115648

A:Accession: PH1525

A:Molecule type: protein

A:Residues: 18-27 <DUP>

A:Experimental source: LAK cell

R:Roiani, M.V.; Finlay, B.B.; Gray, V.; Dedhar, S.

Biochemistry 30, 9859-9866, 1991

A:Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (C

A;Reference number: A40346; MUID:92002034

A;Accession: A40346

A;Molecule type: protein

A;Residues: 18-34; R' <RQ>

R;Krause, K.H.; Simmerman, H.K.B.; Jones, L.R.; Campbell, K.P.

Biochem. J. 270, 545-548, 1990

A;Title: Sequence similarity of calreticulin with a Ca(2+)-binding protein that co-purifies

A;Reference number: S11475; MUID:90380058

A;Accession: S11475

A;Molecule type: protein

A;Residues: 18-32 <KRA>

R;Lamerdin, J.; McCready, P.; Stilwagen, S.; Ramirez, M.; Carrano, A.

submitted to the EMBL Data Library, November 1996

A;Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region

A;Reference number: Z22906

A;Accession: T45075

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-417 <LAMB>

A;Cross-references: EMBL:AD000092; PIDN:AAB51176.1

A;Experimental source: cell line 5HL2-B; fibroblast

C;Comment: Autoantibodies specific for this protein are found in Sjogren's syndrome and

C;Genetics:

A;Gene: CDB:CALR

A;Cross-references: GDB:125179; OMIM:109091

A;Map position: 19p13.3-19p13.2

A;Introns: 31/1; 65/1; 133/1; 164/3; 272/3; 320/3; 351/3

A;Note: CRTG

C;Superfamily: calreticulin

C;Keywords: calcium binding; integrin binding

F;1-17/Domain: signal sequence #status predicted <SIG>

F;18-417/Product: calreticulin #status predicted <MAT>

F;414-417/Region: endoplasmic reticulum retention signal

Query Match 100.0%; Score 2206; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 7.2e-127;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 60
|||
Db 18 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 77
|||
QY 61 SASFEFESNKGQTLVQVQTVKHEQNDICGGYVKLFPSNLDQTMHGDSEYNIMFGPDIC 120
|||
Db 78 SASFEFESNKGQTLVQVQTVKHEQNDICGGYVKLFPSNLDQTMHGDSEYNIMFGPDIC 137
|||
QY 121 GPCTKKVHVIFNFKGNVLINKDIRCKDDEFTHTLTLIVRPDNTYEVKIDNSQVESGSL 180
|||
Db 138 GPCTKKVHVIFNFKGNVLINKDIRCKDDEFTHTLTLIVRPDNTYEVKIDNSQVESGSL 197
|||
QY 181 DWDDELPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDKPEHIPDPAKKPEDWDEEM 240
|||
Db 198 DWDDELPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDKPEHIPDPAKKPEDWDEEM 257
|||
QY 241 DGEWEPVQNPYKGEWKPRQIDNPDYKGTWHPIDNEIDNPEYSPDPSIYAYDNFVGLGLD 300
|||
Db 258 DGEWEPVQNPYKGEWKPRQIDNPDYKGTWHPIDNEIDNPEYSPDPSIYAYDNFVGLGLD 317
|||
QY 301 LWQVKSSTIFDNFLITNDEAYAEFGNETWGTAAKQKQKQKQKQKQKQKQKQKQKQK 360
|||
Db 318 LWQVKSSTIFDNFLITNDEAYAEFGNETWGTAAKQKQKQKQKQKQKQKQKQKQKQK 377
|||
QY 361 EEEAEADKEDDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
|||
Db 378 EEEAEADKEDDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 417
|||
```

RESULT 2

S06763

calreticulin precursor - mouse

N;Alternate names: 55K calcium-binding reticuloplasmin; calregulin

C;Species: Mus musculus (house mouse)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S06763; JCI1444; PC1233; A57498

R;Smith, M.J.; Koch, G.L.E.

EMBO J. 8, 3581-3586, 1989

A;Title: Multiple zones in the sequence of calreticulin. (CRP55, calregulin, HACBP), a

A;Reference number: S06763; MUID:90059955

A;Accession: S06763

A;Molecule type: DNA

A;Residues: 1-416 <SMI>

A;Cross-references: EMBL:X14926; NID:g50567; PIDN:CAA33053.1; PID:g50568

R;Nazzarelli, R.A.; Gold, P.; Cunningham, M.; Green, M.

Gene 120, 217-225, 1992

A;Title: Determination of the sequence of an expressible cDNA clone encoding ERP50/ca

A;Reference number: JCI1444; MUID:93013037

A;Accession: JCI1444

A;Molecule type: mRNA

A;Residues: 1-416 <MAZ>

A;Cross-references: GB:M92988; NID:gi93084; PIDN:AAA37569.1; PID:gi93085

A;Accession: PC1233

A;Molecule type: protein

A;Residues: 18-41 <MA2>

R;White, T.K.; Zhu, Q.; Tanzer, M.L.

J. Biol. Chem. 270, 15926-15929, 1995

A;Title: Cell surface calreticulin is a putative mannoside lectin which triggers mous

A;Reference number: A57498; MUID:95332280

A;Accession: A57498

A;Status: preliminary

A;Molecule type: protein

A;Residues: 74-80;142-151;186-193 <WHI>

C;Superfamily: calreticulin

C;Keywords: calcium binding

F;1-17/Domain: signal sequence #status predicted <SIG>

F;18-416/Product: calregulin #status experimental <MAT>

F;413-416/Region: endoplasmic reticulum retention signal

Query Match 95.2%; Score 2099.5; DB 1; Length 416;
Best Local Similarity 94.2%; Pred. No. 2.1e-120;
Matches 377; Conservative 14; Mismatches 8; Indels 1; Gaps 1;

```
QY 1 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 60
|||
Db 18 DPALYFKEQFLDGDGWTNRWVESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 77
|||
QY 61 SASFEFESNKGQTLVQVQTVKHEQNDICGGYVKLFPSNLDQTMHGDSEYNIMFGPDIC 120
|||
Db 78 SASFEFESNKGQTLVQVQTVKHEQNDICGGYVKLFPSGGLDQKMGDSEYNIMFGPDIC 137
|||
QY 121 GPCTKKVHVIFNFKGNVLINKDIRCKDDEFTHTLTLIVRPDNTYEVKIDNSQVESGSL 180
|||
Db 138 GPCTKKVHVIFNFKGNVLINKDIRCKDDEFTHTLTLIVRPDNTYEVKIDNSQVESGSL 197
|||
QY 181 DWDDELPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDKPEHIPDPAKKPEDWDEEM 240
|||
Db 198 DWDDELPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDKPEHIPDPAKKPEDWDEEM 257
|||
QY 241 DGEWEPVQNPYKGEWKPRQIDNPDYKGTWHPIDNEIDNPEYSPDPSIYAYDNFVGLGLD 300
|||
Db 258 DGEWEPVQNPYKGEWKPRQIDNPDYKGTWHPIDNEIDNPEYSPDANIYAYDNFVGLGLD 317
|||
QY 301 LWQVKSSTIFDNFLITNDEAYAEFGNETWGTAAKQKQKQKQKQKQKQKQKQKQKQK 360
|||
Db 318 LWQVKSSTIFDNFLITNDEAYAEFGNETWGTAAKQKQKQKQKQKQKQKQKQKQKQK 377
|||
QY 361 EEEAEADKEDDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
|||
Db 378 EEEAEADKEDDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 416
|||
```

RESULT 3

JH0819

calreticulin precursor - rat

N;Alternate names: calcium-binding protein 3

A;Residues: 18-29 <TRE>
C;Superfamily: calreticulin
C;Keywords: calcium binding; glycoprotein
F;1-17/DNA/protein: signal sequence #status predicted <SIG>
F;18-416/product: calreticulin #status experimental <MAT>
F;204-412/Region: nuclear location signal
F;413-416/Region: endoplasmic reticulum retention signal
F;344/Binding site: carbonyl group (Asn) (covalent) #status predicted

Query Match 94.8%; Score 2091.5; DB 2; Length 416;
Best Local Similarity 93.8%; Pred. No. 6.3e-120;
Matches 375; Conservative 15; Mismatches 9; Indels 1; Gaps 1;

QY 1 EPAYFRKQFLDGDGWTSRWIESKHSDFKFVLSGGKYGVDEKKDLQTQSODARFYAL 60
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
Db 18 DPATIFRQFLDGDGTNRWVESKHKSDFKFVLSGGKYGVDEKKDLQTQSODARFYAL 77
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
QY 61 SASPEPSNKGQTLVVGTFVKHKEQNIDCGGYVKLPNSLDOTDMHGDSSEYNMFGPDIC 120
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
Db 78 SARPEPSNKGQTLVVGTFVKHKEQNIDCGGYVKLPNGGLDQDMHGDSSEYNMFGPDIC 137
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
QY 121 GPCTKKVHVFNKYGKNVLINKDIRCKDDFTHTLYLIVRPNTYYEKKINQSVESGLE 180
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
Db 138 GPCTKKVHVFNKYGKNVLINKDIRCKDDFTHTLYLIVRPNTYYEKKINQSVESGLE 197
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
QY 181 DWDPLPDKKIOPDAKPEDWDERAKIDPTDSKPEDWKDPHEIPDPDAKPEDWDEM 240
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
Db 198 DWDPLPDKKIOPDAKPEDWDERAKIDPTDSKPEDWKDPHEIPDPDAKPEDWDEM 257
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
QY 241 DGWEPPVIQNPYEKGEMKPRQINDPDYKTWIHPIDNPYESPDPSIYAIDNFVGLGLD 300
:::|:||:||||||| ||:||:||||||| ||:||||||| ||:|||||||
Db 258 DGWEPPVIQNPYEKGEMKPRQINDPDYKTWIHPIDNPYESPDANIYADSAFVGLGLD 317
:::~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
QY 301 LWQVKSITFDNFLITNDLAYAEEFGNETWGVTGAEEKQMOKQODEQRLEKEBEKKRK 360
:::~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
Db 318 LWQVKSITFDNFLITNDLAYAEEFGNETWGVTGAEEKQMOKQODEQRLEKEBEKKRK 377
:::~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
QY 361 EEREAEKDDEDKDEDEPKERDEEDVPQAQDEL 400
|||||||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:
Db 378 EEREAEKDDEDKDEDEPKERDEEDVPQAQDEL 416
|||||||:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:~::~:

RESULT 4
A34154
Calreticulin precursor, skeletal muscle - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Data: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A34154 | S13047
J.R.Fliegel, L.; Burns, K.; MacLennan, D.H.; Reithmeier, R.A.F.; Michalak, M.
J. Biol. Chem. 264, 21522-21528, 1989
A>Title: Molecular cloning of the high affinity calcium-binding protein (calreticulin)
A;Reference number: A34154; MUID:90094320
A;Accession: A34154
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <FLI>
R;Cross-references: GB:J05138; NID:g164858; PIDN:AAA31188.1; PTD:g164859
R;Reviews: S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; M Biochem. J. 271, 473-480, 1990
A>Title: Calreticulin is a candidate for a casequestrin-like function in Ca(2+)-stor
A;Reference number: S13045; MUID:91054414
A;Accession: S13047
A:Molecule type: protein
A:Residues: 19-32 <TR>
C;Superfamily: calreticulin
C;Keywords: skeletal muscle
F;1-17/DNA/protein: signal sequence #status predicted <SIG>
F;415-418/Region: endoplasmic reticulum retention signal

Query Match 93.9%; Score 2070.5; DB 1; Length 418;
Best Local Similarity 91.3%; Pred. No. 1.2e-118.

Matches 374; Conservative 12; Mismatches 14; Indels 1; Gaps 1;

QY 1 EPAYFEKQFLDGGTWSRWIESKHSKDFGKFLVSSGKFGYDDEKDKGLQTSODARFAL 60
 DB 18 EPVYFEKQFLDGGTWSRWIESKHSKDFGKFLVSSGKFGYDDEKDKGLQTSODARFAL 77
 QY 61 SASFEPSNKGQTLVVOFTVKHQNIDCGGYYKFLPNSLDQTMHGDSEYNTMFGPDIC 120
 DB 78 SARFEPSNKGQTLVVOFTVKHQNIDCGGYYKFLPNSLDQTMHGDSEYNTMFGPDIC 137
 QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180
 DB 138 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 197
 QY 181 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDAKKPEDWDEEM 240
 DB 198 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDAKKPEDWDEEM 257
 QY 241 DGEWEPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPENYSPDPSIYAYDNFVGLGLD 300
 DB 258 DGEWEPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPENYSPDPSIYAYDNFVGLGLD 317
 QY 301 LMQVKSCTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKQDEORLKEEEDKKR 360
 DB 318 LMQVKSCTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKQDEORLKEEEDKKR 377
 QY 361 EEEA-EDKEDDKDE 400
 DB 378 EEEA-EDKEDDKDE 418

RESULT 5
 S43376
 calreticulin, brain isoform 1 - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 20-Oct-1994 #sequence_revision 23-Mar-1995 #text_change 07-May-1999
 C:Accession: S43376; S36801
 R:Matsuoka, K.; Seta, K.; Yamakawa, Y.; Okuyama, T.; Shinoda, T.; Isobe, T.
 Biochem. J. 298, 435-442, 1994
 A:Title: Covalent structure of bovine brain calreticulin.
 A:Reference number: S43376; MUID:94183174
 A:Accession: S43376
 A:Molecule type: protein
 A:Residues: 1-400 <MAT>
 A:Experimental source: brain
 R:Liu, N.; Fine, R.E.; Johnson, R.J.
 Biochim. Biophys. Acta 1202, 70-76, 1993
 A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin.
 A:Reference number: S36799; MUID:93385184
 A:Accession: S36801
 A>Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 45-63, 'E', 65-83 <LIU>
 A:Experimental source: brain, clone 8.1
 C:Superfamily: calreticulin
 C:Keywords: calcium binding; glycoprotein
 F:397-400/Region: endoplasmic reticulum retention signal
 F:120-146/Disulfide bonds: #status experimental
 F:162/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 93.3%; Score 2059; DB 2; Length 400;
 Best Local Similarity 93.0%; Pred. No. 5.6e-118;
 Matches 373; Conservative 13; Mismatches 13; Indels 2; Gaps 2;

QY 1 EPAYFEKQFLDGGTWSRWIESKHSKDFGKFLVSSGKFGYDDEKDKGLQTSODARFAL 60
 DB 1 DPTVFEKQFLDGGTWSRWIESKHSKDFGKFLVSSGKFGYDDEKDKGLQTSODARFAL 60
 QY 61 SASFEPSNKGQTLVVOFTVKHQNIDCGGYYKFLPNSLDQTMHGDSEYNTMFGPDIC 120
 DB 61 SARFEPSNKGQTLVVOFTVKHQNIDCGGYYKFLPNSLDQTMHGDSEYNTMFGPDIC 120

QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180
 DB 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180
 QY 181 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDAKKPEDWDEEM 240
 DB 181 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDAKKPEDWDEEM 240
 QY 241 DGEWEPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPENYSPDPSIYAYDNFVGLGLD 300
 DB 241 DGEWEPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPENYSPDPSIYAYDNFVGLGLD 300
 QY 301 LMQVKSCTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKQDEORLKEEEDKKR 360
 DB 301 LMQVKSCTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKQDEORLKEEEDKKR 360
 QY 361 EEEA-EDKEDDKDE 400
 DB 361 EEEA-EDKEDDKDE 400

RESULT 6
 S29129
 calreticulin precursor (clone 3) - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
 C:Accession: S29129
 R:Treves, S.; Zorzato, F.; Pozzan, T.
 Biochem. J. 287, 579-581, 1992
 A:Title: Identification of calreticulin isoforms in the central nervous system.
 A:Reference number: S29129; MUID:93074997
 A:Accession: S29129
 A:Molecule type: mRNA
 A:Residues: 1-411 <TRE>
 A:Cross-references: EMBL:X67597; NID:964608; PIDN:CAA47866.1; PID:964609
 C:Superfamily: calreticulin
 C:Keywords: glycoprotein
 F:1-12/Domain: signal sequence (fragment) #status predicted <Sig>
 F:13-411/Product: calreticulin #status predicted <MAT>
 F:408-411/Region: endoplasmic reticulum retention signal
 F:339/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.3%; Score 1837.5; DB 2; Length 411;
 Best Local Similarity 80.5%; Pred. No. 1.6e-104;
 Matches 323; Conservative 44; Mismatches 31; Indels 3; Gaps 2;

QY 1 EPAYFEKQFLDGGTWSRWIESKHSKDFGKFLVSSGKFGYDDEKDKGLQTSODARFAL 60
 DB 13 EPAYFEKQFLDGGTWSRWIESKHSKDFGKFLVSSGKFGYDDEKDKGLQTSODARFAL 72
 QY 61 SASFEPSNKGQTLVVOFTVKHQNIDCGGYYKFLPNSLDQTMHGDSEYNTMFGPDIC 120
 DB 73 SRFDSFSNKKDQTLVVOFVSVKHQNIDCGGYYKFLPNSLDQTMHGDSEYNTMFGPDIC 132
 QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180
 DB 133 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 192
 QY 181 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDAKKPEDWDEEM 240
 DB 193 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDAKKPEDWDEEM 252
 QY 241 DGEWEPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPENYSPDPSIYAYDNFVGLGLD 300
 DB 253 DGEWEPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPENYSPDPSIYAYDNFVGLGLD 312
 QY 301 LMQVKSCTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKQDEORLKEEEDKKR 360
 DB 313 LMQVKSCTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKQDEORLKEEEDKKR 372
 QY 361 EEEA-EDKEDDKDE 400
 DB 361 EEEA-EDKEDDKDE 400

2

F;418-421/Region: endoplasmic reticulum retention signal				
Query Match	54.4%;	Score 1201;	DB 2;	Length 421;
Best.Local Similarity	54.8%;	Pred. No. 6	6e-66;	
Matches 223;	Conservative 67;	Mismatches 95;	Indels 22;	Gaps
QY	4	VYFKEQFLDGGWTSRWIESKHKSD----	FGKFVLSSGKFYGDDEKDKGLQTSODARFYAL	60
Db	27	VFFOEKF--EDGWESRWKSEWKAKDENMAGEWNHTSGKNGDAE--DKGIQTSEIDYRFYAI	83	
QY	61	SASFEPESNKGQTLIVQVTFVKVHEQNTDCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC	120	
Db	84	SABYPEFSNKDKTLVLQFSYKHEQKLDCCGGYVKLLGGVDQKFGDTSYSIMFGPDIC	143	
QY	121	GPGTGKVHVIFNFKGKNVLINKDKIRCKDDEFTLHLYTLIVRPDNTYEVKIDNSOVESGSLE	180	
Db	144	GYSTKKVHTTILTDGKNHLLIKDVPCETDQTLTHVYTLIIRPDATYSILIDNEEKQGTGSIY	203	
QY	181	DDMDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKSPEDWDK--PEHIPDPDAKKPEDWDEE	239	
Db	204	EHMDILPPKKIKDPEAKKPEDDDQKEYIPDPEKKPEGYDDIPKEIPDPDAKKPEDWDEE	263	
QY	240	MDGEWEPVPTQNEPYKGEWKPRQIDNPDPYKGTWTHPEIDNPEYSPDPSIYAYDNFGVIGL	299	
Db	264	EDGEWTAPTIPNPEYKGPWKQKKIKNPYOGKKWAPMIDNPDKDPPIYAFDLSKYIGI	323	
QY	300	DLQVKSGLTFIDNPLTNDAYAAEEFGNETWGTVTKAAEQMKQDKQDEEQRLKEEEDDKR	359	
Db	324	ELQVKSGLTFDNIITDDPALAKTFAEETWGWKHKEAKAAAFDAEK----KKEEEDAOK	379	
QY	360	---KKEEEAEDKDEDEKDEDEED---EDKDEDEDEBVPQAKDEL	400	
Db	380	GGDEDDDLDEDEDEKADKADSADSDSKSDSDSKQ-----HDEL	421	

Search completed: October 21, 2001, 03:07:44
Job time: 339 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:03:50 ; Search time 28.98 Seconds
(without alignments)
472.816 Million cell updates/sec

Title: US-09-807-148-3
Perfect score: 2206
Sequence: 1 EPAYVFKQFLDGDGWTSRW.....EEDKEEDBEDVPGQAKDEL 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2206	100.0	417	1	CRTC_HUMAN
2	2099.5	95.2	416	1	P27797 homo sapien
3	2091.5	94.8	416	1	P14211 mus musculus
4	2070.5	93.9	418	1	P15253 oryctolagus
5	2059	93.3	400	1	P52193 bos taurus
6	1641.5	74.4	421	1	P42918 bos taurus
7	1461.5	66.3	406	1	P29413 drosophila
8	1400.5	63.5	395	1	P27798 caenorhabdi
9	1366.5	61.9	388	1	P11012 onchocerca
10	1206.5	54.7	415	1	P93508 ricinus com
11	1171.5	53.1	416	1	Q40401 nicotiana p
12	1169.5	53.0	421	1	Q9xf98 prunus arme
13	1165	52.8	424	1	Q38858 arabidopsis
14	1164.5	52.8	401	1	Q92ny3 euglena gra
15	1161.5	52.7	416	1	O81919 beta vulgar
16	1161.5	52.7	420	1	Q9sp22 zea mays (m
17	1150.5	52.2	420	1	Q9std3 chlamydomon
18	1150.5	52.2	425	1	O04151 arabidopsis
19	1145.5	51.9	416	1	O9zpp1 berberis st
20	1135	51.5	424	1	O9slv8 oryza sativ
21	1123.5	50.9	424	1	O23858 dictyosteli
22	1094.5	49.6	393	1	O06814 schistosoma
23	989	44.8	424	1	O04153 arabidopsis
24	810	36.7	591	1	P35564 mus musculus
25	803	36.4	591	1	P35565 rattus norv
26	799	36.2	592	1	P27824 homo sapien
27	799	36.2	593	1	P24643 canis famil
28	731.5	33.2	610	1	O14967 homo sapien
29	718	32.5	611	1	P52194 mus musculus
30	676	30.6	546	1	Q39817 glycine max
31	658	29.8	530	1	P29402 arabidopsis
32	631.5	28.6	551	1	O82709 pisum sativ
33	628.5	28.5	560	1	P36581 schizosacch

34 618.5 28.0 619 1 CALX_CAEEL
35 608 27.6 528 1 CAX2_ARATH
36 590 26.7 540 1 CALX_HELTU
37 447 20.3 105 1 CRTC_PIG
38 385.5 17.5 502 1 CALX_YEAST
39 186.5 8.5 411 1 MP62_LYTP1
40 184.5 8.4 713 1 NUCLE_MESAU
41 183.5 8.3 743 1 ABRA_PLAFC
42 177.5 8.0 706 1 NUCLE_MOUSE
43 176 8.0 694 1 NUCLE_CHICK
44 175.5 8.0 712 1 NUCLE_RAT
45 173.5 7.9 321 1 ABRA_PLAFF

ALIGNMENTS

RESULT 1

CRTC_HUMAN
ID CRTC_HUMAN STANDARD; PRT; 417 AA.
AC P27797;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (52 KDA
DE RIBONUCLEOPROTEIN AUTOANTIGEN RO/SS-A).
GN CALR OR CRTC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92013129; PubMed=1919005;
RA Rokeach L.A., Haselby J.A., Meilof J.F., Smeenk R.J., Unnasch T.R.,
RA Greene B.M., Hoch S.O.;
RT "Characterization of the autoantigen calreticulin.";
RL J. Immunol. 147:3031-3039(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90237213; PubMed=2332496;
RA McCauliffe D.P., Lux F.A., Lieu T.S., Sanz I., Hanke J., Newkirk M.M.,
RA Bachinski L.L., Itoh Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,
RA Capra J.D.;
RT "Molecular cloning, expression, and chromosome 19 localization of a
human RO/SS-A autoantigen.";
RL J. Clin. Invest. 85:1379-1391(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92129342; PubMed=1733953;
RA McCauliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;
RT "The 5'-flanking region of the human calreticulin gene shares
RT homology with the human GRP78, GRP94, and protein disulfide isomerase
RT promoters.";
RL J. Biol. Chem. 267:2557-2562(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA Lamerdin J., McCreedy P., Stilwagen S., Ramirez M., Carrano A.;
RT "Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases."
RN [5]
RP SEQUENCE OF 18-36.
RX MEDLINE=92020234; PubMed=1911778;
RA Rojiani M.V., Finlay B.B., Gray V., Dedhar S.;
RT "In vitro interaction of a polypeptide homologous to human RO/SS-A
RT antigen (calreticulin) with a highly conserved amino acid sequence in
RT the cytoplasmic domain of integrin alpha subunits.";
RL Biochemistry 30:9859-9866(1991).
RN [6]
RP SEQUENCE OF 18-32.
RX MEDLINE=90380058; PubMed=2400400;
RA Ktause K.H., Simmerman H.K.B., Jones L.R., Campbell K.P.;
RT "Sequence similarity of calreticulin with a Ca2(+)-binding protein
that co-purifies with an Ins(1,4,5)P3-sensitive Ca2+ store in HL-60

cells."; Biochem. J. 270:545-548(1990).
 (7)
 RN SEQUENCE OF 18-28.
 RP TISSUE=Liver;
 RC MEDLINE=93162045; PubMed=1286669;
 RX Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
 RA Appel R.D., Hughes G.J.;
 RT "Human liver protein map: a reference database established by
 RT microsequencing and gel comparison.";
 RL Electrophoresis 13:992-1001(1992).
 RN [8]
 RN PARTIAL SEQUENCE OF 25-34; 56-62; 208-221 AND 273-278.
 RP TISSUE=Keratinocytes;
 RC MEDLINE=93162043; PubMed=1286667;
 RX Rasmussen H.H., van Damme J., Puype M., Gesser B., Cellis J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [9]
 RN SEQUENCE OF 18-26.
 RP TISSUE=Colon carcinoma;
 RC MEDLINE=97295306; PubMed=9150948;
 RX Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
 RA "A two-dimensional gel database of human colon carcinoma proteins.";
 RL Electrophoresis 18:605-613(1997).
 CC !- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC !- LOW AFFINITY CALCIUM-BINDING SITES.
 CC !- SUBUNIT: MONOMER (BY SIMILARITY).
 CC !- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC !- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC !- CAUTION: WAS ORIGINALLY (REF 2) THOUGHT TO BE THE RO AUTOANTIGEN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M84739; AAA51916.1; -;
 DR EMBL; M32294; AAA36582.1; -;
 DR EMBL; AD000092; AAB51176.1; -;
 DR PIR; A37047; A37047.
 DR PIR; S11475; S11475.
 DR PIR; A42330; A42330.
 DR PIR; A46452; A46452.
 DR SWISS-2DPAGE; P27797; HUMAN.
 DR Aarhu/Chent-2DPAGE; 9401; IEF.
 DR HSC-2DPAGE; P27797; HUMAN.
 DR MIM; 109091; -;
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 417 CALRETICULIN.
 FT DOMAIN 18 197 N-DOMAIN.
 FT DOMAIN 198 308 P-DOMAIN.
 FT DOMAIN 309 417 C-DOMAIN.
 FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
 FT REPEAT 191 202 1-1.
 FT REPEAT 210 221 1-2.
 FT REPEAT 227 238 1-3.
 FT REPEAT 244 255 1-4.

FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
 FT REPEAT 259 269 2-1.
 FT REPEAT 273 283 2-2.
 FT REPEAT 287 297 2-3.
 FT DOMAIN 351 408 ASP/GLU/LYS-RICH.
 FT DISULFID 137 163 BY SIMILARITY.
 FT SITE 414 417 PREVENT SECRETION FROM ER.
 FT CONFLICT 35 35 MISSING (IN REF. 3).
 SQ SEQUENCE 417 AA; 48141 MW; BC37C3C0F1054FB2 CRC64;
 .
 .
 Query Match 100.0%; Score 2206; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1e-119;
 Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1. EPAVYFKEQFLDGDGWTSRWIESKHKSDFKFLVSSGKFGYDGEKDKGLQTSQDARFYAL 60
 DB 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFKFLVSSGKFGYDGEKDKGLQTSQDARFYAL 77
 QY 61 SASFEFPSNKGQTLVVOFTVKHEQNTDCGGYVKLPNSLDQTDMDHGDSEYNIMFGPDIC 120
 DB 78 SASFEFPSNKGQTLVVOFTVKHEQNTDCGGYVKLPNSLDQTDMDHGDSEYNIMFGPDIC 137
 QY 121 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
 DB 138 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 197
 QY 181 DWDWFLPPKKIKDPASKPDMDERAKIDDPDTSKDEWDKPEHIIPDPAKPEDWDEEM 240
 DB 198 DWDWFLPPKKIKDPASKPDMDERAKIDDPDTSKDEWDKPEHIIPDPAKPEDWDEEM 257
 QY 241 DGEWEPVIONPEYKGEWKPQINDPDYKGTWTHPEIDNPEYSPDPSIAYDNFVGLGLD 300
 DB 258 DGEWEPVIONPEYKGEWKPQINDPDYKGTWTHPEIDNPEYSPDPSIAYDNFVGLGLD 317
 QY 301 LMQVKSCTIFDNFLINDAYAEFGNETGWVTKAAEKQMKQDDEQRLEKEEDKKRK 360
 DB 318 LMQVKSCTIFDNFLINDAYAEFGNETGWVTKAAEKQMKQDDEQRLEKEEDKKRK 377
 QY 361 EREAEADKEDDEKDE 400
 DB 378 EREAEADKEDDEKDE 417
 RESULT 2
 CRTX_MOUSE
 ID CRTX_MOUSE STANDARD; PRT; 416 AA.
 AC P14211;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (BRP60).
 GN CALR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-48 AND 129-161.
 RC STRAIN-BALB/C. TISSUE=Liver;
 RX MEDLINE=90059955; PubMed=2583110;
 RA Smith M.J., Koch G.L.E.;
 RT "Multiple zones in the sequence of calreticulin (CRP55, calregulin,
 RT HACBP), a major calcium binding ER/SR protein.";
 RL EMBO J. 8:3581-3586(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93013037; PubMed=1398135;
 RA Mazzarella R.A., Gold P., Cunningham M., Green M.;
 RT "Determination of the sequence of an expressible cDNA clone encoding
 RT ERp60/calregulin by the use of a novel nested set method.";
 RL Gene 120:217-225(1992).
 RN [3]

SEQUENCE OF 18-38.
 CC TISSUE-Fibroblast;
 RA MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RL using preparative two-dimensional gel electrophoresis.";
 CC Electrophoresis 15:735-745(1994).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X14926; CAA33053.1; -;
 CC EMBL; M92988; AAA37569.1; -;
 CC PIR; S06763; S06763.
 CC PIR; JC1444; JC1444.
 CC SWISS-2DPAGE; P14211; MOUSE.
 CC MGI; MGI:88252; Calr.
 CC InterPro; IPR000886; -;
 CC InterPro; IPR001580; -;
 CC Pfam; PF00262; calreticulin; 1.
 CC PRINTS; PR00626; calreticulin;
 CC PROSITE; PS00014; ER_TARGET; 1.
 CC PROSITE; PS00803; CALRETICULIN_1; 1.
 CC PROSITE; PS00804; CALRETICULIN_2; 1.
 CC PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 CC Endoplasmic reticulum; Calcium-binding; Repeat; signal.
 FT SIGNAL 1 17
 FT CHAIN 18 416 CALRETICULIN.
 FT DOMAIN 18 197 N-DOMAIN.
 FT DOMAIN 198 308 P-DOMAIN.
 FT DOMAIN 309 416 C-DOMAIN.
 FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
 FT REPEAT 191 202 1-1.
 FT REPEAT 210 221 1-2.
 FT REPEAT 227 238 1-3.
 FT REPEAT 244 255 1-4.
 FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
 FT REPEAT 273 283 2-1.
 FT REPEAT 287 297 2-2.
 FT REPEAT 287 297 2-3.
 FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
 FT DISULFID 137 163 BY SIMILARITY.
 FT SITE 413 416 PREVENT SECRETION FROM ER.
 SQ SEQUENCE 416 AA; 47994 MW; 24C03B00913408D8 CRC64;

Query Match 95.28; Score 2099.5; DB 1; Length 416;
 Best Local Similarity 94.28; Pred. No. 1.3e-113;
 Matches 377; Conservative 14; Mismatches 8; Indels 1; Gaps 1;
 Qy 1 EPAYVEKQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYGDGKGLQTSQDARFYAL 60
 Db 18 DPAYIFKRFQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYGDGKGLQTSQDARFYAL 77
 Qy 61 SASFPFNSKGGTLLVQFTVAKHEQNCGGYVKLFPPNSLQDTHGDSYNIMFGPDIC 120
 Db 78 SAKFEPFNSKGGTLLVQFTVAKHEQNCGGYVKLFPPNSLQDTHGDSYNIMFGPDIC 137
 Qy 121 GPGTKKVVHVFYKGNVLYNKIRCKDDETHLYTLVVRPNTYEVKIDNSQVSGSLE 180
 Db 138 GPGTKKVVHVFYKGNVLYNKIRCKDDETHLYTLVVRPNTYEVKIDNSQVSGSLE 197
 Qy 181 DDWFLPFPKKIKDPDASKPDPEDWDERAKIDDTDSKPEDWDKPEHIPDPDAKKPEDWDEEM 240

Db 198 DDWFLPFPKKIKDPDASKPDPEDWDERAKIDDTDSKPEDWDKPEHIPDPDAKKPEDWDEEM 257
 Qy 241 DGEWPPVIQNPYKGEWKPRQIDNPDKYKGIWIHPEIDNPYSPDPSPYIAYDNFVGLGD 300
 Db 233 DGEWPPVIQNPYKGEWKPRQIDNPDKYKGIWIHPEIDNPYSPDPSPYIAYDNFVGLGD 317
 Qy 301 LQWVKSQTIFDNFLITNDENAYAEFGNETGWTKAAEKQMKDKQDEORLKEEEDKKRK 360
 Db 318 LQWVKSQTIFDNFLITNDENAYAEFGNETGWTKAAEKQMKDKQDEORLKEEEDKKRK 377
 Qy 361 EEEAEKDEDEDKDE 400
 Db 378 EEEAEKDEDEDKDE 416
 RESULT 3
 CRQC_RAT
 ID CRQC_RAT STANDARD; PRT; 416 AA.
 AC P18418; P10452;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (CALBP)
 DE (CALCIUM-BINDING PROTEIN 3) (CABP3).
 GN CALR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain cortex;
 RX MEDLINE=90370496; PubMed=2395661;
 RA Murthy K.K., Banville D., Srikanth C.B., Carrier F., Bell A.,
 RA Holmes C., Patel Y.C.;
 RT "Structural homology between the rat calreticulin gene product and
 RT the Onchocerca volvulus antigen Rai-1.";
 RL Nucleic Acids Res. 18:4933-4933(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=93202172; PubMed=8453984;
 RA Nakamura M., Moriya M., Baba T., Michikawa Y., Yamanobe T., Arai K.,
 RA Okinaga S., Kobayashi T.;
 RT "An endoplasmic reticulum protein, calreticulin, is transported into
 RT the acrosome of rat sperm.";
 RL Exp. Cell Res. 205:101-110(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
 RX MEDLINE=95181573; PubMed=7876339;
 RA Soennichsen B., Fueflekrug J., van Nguyen P., Diekmann W.,
 RA Robinson D.G., Mieskes G.;
 RT "Retention and retrieval: both mechanisms cooperate to maintain
 RT calreticulin in the endoplasmic reticulum.";
 RL J. Cell Sci. 107:2705-2717(1994).
 RN [4]
 RP SEQUENCE OF 270-358 FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX Leona Y.C., Bailly A., Latruffe N.;
 RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 18-29.
 RX MEDLINE=91054414; PubMed=2241926;
 RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,
 RA MacLennan D.H., Meldolesi J., Pozzan T.;
 RT "Calreticulin is a candidate for a calsequestrin-like function in
 RT Ca2+-storage compartments (calcosomes) of liver and brain.";
 RL Biochem. J. 271:473-480(1990).
 RN [6]
 RP SEQUENCE OF 18-32.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;

RX MEDLINE-92360010; PubMed-1497655;
RA Nakamura M., Michikawa Y., Baba T., Okinaga S., Arai K.;
RT "Calreticulin is present in the acrosome of spermatids of rat
testis.";
RL Biochem. Biophys. Res. Commun. 186:668-673(1992).
RN [7].
RP SEQUENCE OF 18-32.
RC STRAIN=LEC; TISSUE=Liver;
RX MEDLINE-94072621; PubMed-8251535;
RA Yokoi T., Nagayama S., Kajiwara R., Kawaguchi Y., Horiuchi R.,
RT "Identification of protein disulfide isomerase and calreticulin as
autoimmune antigens in LEC strain of rats.";
RL Biochim. Biophys. Acta 1158:339-344(1993).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE D-BETA-
HYDROXYBUTYRATE DEHYDROGENASE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; D78308; BAA11345.1; -;
DR EMBL; X53363; CAA37446.1; -;
DR EMBL; X13702; CAA31987.1; ALT_SEQ.
DR EMBL; X79327; CAA55890.1; -;
DR PIR; S04867; S04867.
DR PIR; S11205; S11205.
DR PIR; S13045; S13045.
DR PIR; A49176; A49176.
DR PIR; S45036; S45036.
DR PIR; JH0819; JH0819.
DR InterPro; IPR000886; -;
DR InterPro; IPR001580; -;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PRO0626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN.1; 1.
DR PROSITE; PS00804; CALRETICULIN.2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1
FT CHAIN 17
FT CHAIN 18 416 CALRETICULIN.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 416 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT REPEAT 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 413 416 PREVENT SECRETION FROM ER.
FT SEQUENCE 416 AA; 47995 MW; 2E6713CED31A2970 CRC64;

Query Match 94.8%; Score 2091.5; DB 1; Length 416;
Best Local Similarity 93.8%; Pred. No. 3.6e-113;
Matches 375; Conservative 15; Mismatches 9; Indels 1; Gaps 1;

QY 1 EPAYFKEQELDGGWTSRWIESKHSDFKFKVLSGKFGYDEKDKGLQTSQDARFAL 60
DB DPALYFKEQELDGGWTSRWIESKHSDFKFKVLSGKFGYDQEKDKGLQTSQDARFAL 77
QY 61 SASFEPSNKGOTLVQFTVKHQNIDCGGYVKLFNSLDQTDHMGDSEYNMFGPDIC 120
DB SARFEPSNKGOTLVQFTVKHQNIDCGGYVKLFPGGLDQDKMHGDSEYNMFGPDIC 137
QY 121 GPGTKVHVIFNYKGNVLINKIDIRCKDDFTHTLYLIVRPDNTYEVKIDNSQVSGSLE 180
DB GPGTKVHVIFNYKGNVLINKIDIRCKDDFTHTLYLIVRPDNTYEVKIDNSQVSGSLE 197
QY 181 DDMDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPDPDAKKPEDMDEEM 240
DB DDMDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPDPDAKKPEDMDEEM 257
QY 241 DGWEPPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPESYDPDSIYAYDNFVGLGLD 300
DB DGWEPPVIONPEYKGEWKPRQIDNPDKGTWTHPEIDNPESYDPDSIYAYDNFVGLGLD 317
QY 301 LWQVKSCTIFDNFLITNDEAYAEFGNETGWVTKAAEKQMKDKODEQRLKEEEDKKRK 360
DB LWQVKSCTIFDNFLITNDEAYAEFGNETGWVTKAAEKQMKDKODEQRLKEEEDKKRK 377
QY 361 EEEAEADKEDDEKDE 400
DB EEEAEADKEDDE 416

RESULT 4
CRTC_RABIT
ID CRTC_RABIT STANDARD; PRT; 418 AA.
AC P15253;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60).
GN CALR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Slow-twitch skeletal muscle;
RX MEDLINE-90094320; PubMed-2600080;
RA Fliegel L., Burns K., MacLennan D.H., Reithmeier R.A.F., Michalak M.;
RT "Molecular cloning of the high affinity calcium-binding protein
(calreticulin) of skeletal muscle sarcoplasmic reticulum.";
RL J. Biol. Chem. 264:21522-21528(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fast-twitch skeletal muscle;
RX MEDLINE-91282795; PubMed-2059224;
RA Fliegel L., Michalak M.;
RT "Fast-twitch and slow-twitch skeletal muscles express the same
isoform of calreticulin.";
RL Biochem. Biophys. Res. Commun. 177:979-984(1991).
RN [3]
RP SEQUENCE OF 18-36.
RX MEDLINE-91054414; PubMed-2241926;
RA Treves S., de Mattei M., Lanfretti M., Villa A., Green N.M.,
RA MacLennan D.H., Meldolesi J., Pozzan T.;
RT "Calreticulin is a candidate for a calsequestrin-like function in
Ca2(+)-storage compartments (calciosomes) of liver and brain.";
RL Biochem. J. 271:473-480(1990).
RN [4]
RP SEQUENCE OF 18-46.
RX MEDLINE-91201375; PubMed-2016321;
RA Milner R.E., Baksh S., Shemanko C., Carpenter M.R., Smillie L.,
RA Vance J.E., Opas M., Michalak M.;
RT "Calreticulin, and not calsequestrin, is the major calcium binding
protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic

	78	SARFEPFSNKGQPLVVOFTVKVHEQNIDCGGGYVKLPAGLDQKMDHGDSYEINIMFGPDIC	13
Db			
QY	.121	GPCTKKVHVLFNTYKGNVLINKDIRCKDDTFHLYTLIVRPDNTYEVKIDNSOVESGSLE	180
Db	.138	GPCTKKVHVLFNTYKGNVLINKDIRCKDDTFHLYTLIVRPDNTYEVKIDNSOVESGSLE	197
QY	.181	DWDNFPPKKIKDPDASKPEDWDERAKIIDPDTOSKPEDMDKPEHIIPDPAKPKEDWDEEM	240
Db	.198	DWDNFPPKKIKDPDASKPEDWDERAKIIDPDTOSKPEDMDKPEHIIPDPAKPKEDWDEEM	257
QY	.241	DGEWEPPVIQNPYKGEWKPRQIDNDPYKGTWIHPETIDNPSPDISIAYDNFVGVLGD	300
Db	.258	DGEWEPPVIQNPYKGEWKPRQIDNDPYKGTWIHPETIDNPSPDIAYDSFAVLGLD	317
QY	.301	LWOVKSGTIFDNFLITNDEAYAEFGNETWGVTAAEKOMKMQRKDQEORLKEEEDKKRK	360
Db	.318	LWOVKSGTIFDNFLITNDEAYAEFGNETWGVTAAEKOMKMQRKDQEORLKEEEDKKRK	377
QY	.361	EFEEA-EDKEDDEDKDEDEDEKDEEDEEDVPGAKDEL	400
		: ::: :	
Db	.378	EFEAEAERDEDKDDEDEDEKDEDEEEAAGAQAQDEL	418
<hr/>			
	RESULT	5	
	CRTL_BOVIN	STANDARD;	PRT; 400 AA.
ID	CRTL_BOVIN		
AC	P52193;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	CALRETICULIN, BRAIN ISOFORM 1 (CRP55) (CALREGULIN) (HACBP).		
OS	Bos taurus (Bovine)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;		
OC	Bovidae; Bovinae; Bos.		
ON	NCBI_Taxid=9913;		
RX	[1]		
RP	SEQUENCE.		
RC	TISSUE=Brain;		
RX	MEDLINE=94183174; PubMed=8135753;		
RA	Matsuoka K., Seta K., Yamakawa Y., Okuyama T., Shinoda T., Isobe T.;		
RL	"Covalent structure of bovine brain calreticulin.";		
RL	Biochem. J. 298:435-442(1994).		
CC	-!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND LOW AFFINITY CALCIUM-BINDING SITES.		
CC	-!- SUBUNIT: MONOMER.		
CC	-!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.		
CC	-!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.		
DR	InterPro: IPR001580; "		
DR	Pfam: PF00262; Calreticulin; 1.		
DR	PRINTS: PR00626; CALRETICULIN.		
DR	PROSITE: PS00014; ER_TARGET; 1.		
DR	PROSITE: PS00803; CALRETICULIN_1; 1.		
DR	PROSITE: PS00804; CALRETICULIN_2; 1.		
DR	PROSITE: PS00805; CALRETICULIN_REPEAT; 3.		
KW	Endoplasmic reticulum; Calcium-binding; repeat; Glycoprotein.		
FT	DOMAIN 1 180 N-DOMAIN.		
FT	DOMAIN 181 291 P-DOMAIN.		
FT	DOMAIN 292 400 C-DOMAIN.		
FT	DOMAIN 174 238 4 X APPROXIMATE REPEATS.		
FT	REPEAT 174 185 1-1.		
FT	REPEAT 193 204 1-2.		
FT	REPEAT 210 221 1-3.		
FT	REPEAT 227 238 1-4.		
FT	DOMAIN 242 280 3 X APPROXIMATE REPEATS.		
FT	REPEAT 242 252 2-1.		
FT	REPEAT 256 266 2-2.		
FT	REPEAT ⁴ 270 280 2-3.		
FT	DOMAIN 334 390 ASP/GLU/LYS-RICH.		
FT	DISULFID 120 146		
FT	CARBOHYD 162 162		
FT	SITE 397 400 N-LINKED (GLCNAC. .). PREVENT SEQUENCE FROM ER (POTENTIAL).		

[illegible]

OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93208374; PubMed=1296819;
 RA Smith M.J.;
 RT "Nucleotide sequence of a Drosophila melanogaster gene encoding a
 RT calreticulin homologue.";
 RL DNA Seq. 3:247-250(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anandakrishnan P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson M.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacht J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrlkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 91-124 AND 182-220.
 RX MEDLINE=90307981; PubMed=2365822;
 RA McCaulliffe D.P., Zappi E., Lieu T.S., Michalak M., Sontheimer R.D.,
 RA Capra J.D.;
 RT "A human Ro/SS-A autoantigen is the homologue of calreticulin and is
 RT highly homologous with onchocercal RAL-1 antigen and an alysia
 RT 'memory molecule'";
 RL J. Clin. Invest. 86:332-335(1990).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL; X64461; CAA45791.1; -;
 DR EMBL; AE003683; AAF54416.1; -;
 DR PIR; A37158; A37158;
 DR FlyBase; FBgn000585; Crc.
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 406 CALRETICULIN.
 FT CONFLICT 107 107 G -> A (IN REF. 3).
 FT CONFLICT 184 184 V -> L (IN REF. 3).
 FT CONFLICT 184 184 V -> L (IN REF. 3).
 SQ SEQUENCE 406 AA; 46808 MW; 65D72C69D0BEC427 CRC64;

 Query Match 66.3%; Score 1461.5; DB 1; Length 406;
 Best Local Similarity 68.0%; Pred. No. 3.7e-77;
 Matches 263; Conservative 45; Mismatches 76; Indels 3; Gaps 3;
 QY . 4 VYFKEQFLDGCWTSWIESKHK - SDFGKFLVSSGKFGYDEKDKGLQTSODARFYALSA 62
 DB 21 VYLKENF-DNENWEDTWIYKPKGKFGKFLTEGTFTYNDAAEADKGIQTSODARFYAASR 79
 QY 63 SFEPFSNKGQTLVYQVFTVKHSEONDCGGYVKLFPSNLDQTDHMGDSYNTMFGPDICGP 122
 DB 80 KFDGFSNEDKPLVQVFSVKEHQNDICGGYVKLFDCSLDQTDHMGESYETMFGPDICGP 139
 QY 123 GTKKVHVIFNYKGNVLINKDIRCKDDDETHLYTLIVRPDNTYEVKIDNSQVSGSLEDD 182
 DB 140 GTKKVHVIFSYKGNHLISKDIRCKDDVYTFYTLIVRPDNTYEVKIDNSQVSGSLEDD 199
 QY 183 WDLPPKKIKDPDASKPEDWDERAKIDPTDSKPEDWDPKPEHIPPDAKKPEDWDEMDG 242
 DB 200 WDFLAPKKIKDPATKPEDWDRATIPDDPKPEDWDPKPEHIPPDAKKPEDWDEMDG 259
 QY 243 EWEPVPTQNPYKGEWPKQIDNPYKGTWHPETDNPYSPDPSIAYDNFVGLGLDW 302
 DB 260 EWEPMDIDNPEKGEWPKQIDNPYKGTWHPETDNPYSPDPSIAYDNFVGLGLDW 319
 QY 303 QVKSQTIDNPLINDEAYAEFGNETWGTAKAAQKQKQKQKQKQKQKQKQKQKQKQKQK 362
 DB 320 QVKSQTIDNPLINDEAYAEFGNETWGTAKAAQKQKQKQKQKQKQKQKQKQKQKQKQK 378
 QY 363 EEAEKDEDEKDE 389
 DB 379 DDE 405

 RESULT 8
 CRTC_CAEEL STANDARD; PRT; 395 AA.
 ID CRTC_CAEEL
 AC P27798;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE 01-FEB-1994 (Rel. 28, Last annotation update)
 DE CALRETICULIN PRECURSOR.
 GN CRTC1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=92329978; PubMed=1627827;
 RA Smith M.J.;
 RT "A C. elegans gene encodes a protein homologous to mammalian

RT calreticulin.";
RL DNA Seq. 2:235-240(1992).
CC !- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC !- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC !- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59589; CAA42159.1; -
DR PIR: S25851; S25851; -
DR InterPro: IPR000886; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00014; ER.TARGET.1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN ? 395 CALRETICULIN.
FT DOMAIN ? 192 N-DOMAIN.
FT DOMAIN 193 301 P-DOMAIN.
FT DOMAIN 302 395 C-DOMAIN.
FT DOMAIN 186 250 4 X APPROXIMATE REPEATS.
FT REPEAT 186 197 1-1.
FT REPEAT 205 216 1-2.
FT REPEAT 222 233 1-3.
FT REPEAT 239 250 1-4.
FT DOMAIN 254 292 3 X APPROXIMATE REPEATS.
FT REPEAT 254 264 2-1.
FT REPEAT 268 278 2-2.
FT REPEAT 282 292 2-3.
FT DOMAIN 332 390 ASP/GLU/LYS-RICH.
FT DISULFID 133 158 BY SIMILARITY.
FT SITE 392 395 PREVENT SECRETION FROM ER.
SQ SEQUENCE 395 AA; 45616 MW; 35CA7D2EC1D56B03 CRC64;

Query Match 63.5%; Score 1400.5; DB 1; Length 395;
Best Local Similarity 64.7%; Pred. No. 1.1e-73; Mismatches 77; Indels 3; Gaps 3;
Matches 246; Conservative 54;

QY 4 VYFKEQFDGDDGWTSTWIESKHKSDFGKFLVSSGKFGYDGEKDKGLQTSQDARFALSAS 63
Db 17 VYFKEEFDA-SWEKRWQSKHKDDFGAFKLSGKFFDVESRDQGIQTSDAKFYFRAAK 75

QY 64 FE-PFSNKGOTLVQVOTVTHQENIDCGGVYKFLPNSLDQTMHGDSSEYNIMFGPDICGP 122
Db 76 FDKDFSNKGLTVIQYTVRHEQIDCGGVYKVMRADADLDFGHGETPYNVMFGPDICGP 135

QY 123 GTRKVVHVIYKGNVLIINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSLEDD 182
Db 136 -TRRVHVIILYKGENILIKKEITCKSDLETHLYTLFNSDNTYEVKIDGESAQTSLEED 194

QY 183 WDFLPKKIKDPDASKPDWDERAKIDDDTDSKPDWDKPEHIDPDPAKPEDWDEMDG 242
Db 195 WDLPPAKKIKDPDASKPDWDEREYIDDAEDAKPEDWEPKPEHIDPDPAKPEDWDEMDG 254

QY 243 EWEPPVQNPYKGEWKPKQIDNPDPYKGIWTHIHPIDNPYSPDPSIYAYDNFGVIGLDLW 302
Db 255 EWEPPMIDNPYKGEWKPKQIKNPAYKGIWTHIHPIDNPYSPDPSIYAYDNFGVIGLDLW 314

QY 303 QVKSQGTIFDNELITNDVAEAEFGNETWGTVAKEQKMDKDEORLKEEEDKKREE 362
Db 315 QVKSQGTIFDNIIITDSVEAEAAHAAETFDKLTVEKKEKKADEETRKAEERKAESE 374

QY 363 EBAEDKEDDEKDEDEDEDE 382
Db 375 KEAKRDDDEEKEEKEEGHDE 394

RESULT. 9
RALL ONCVO STANDARD; PRT; 388 AA.
ID RALL ONCVO STANDARD; PRT; 388 AA.
AC R11012;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RAL-1 PROTEIN PRECURSOR (41 KDA LARVAL ANTIGEN).
GN RAL1.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94341871; PubMed=7520419;
RA Rokeach L.A., Zimmerman P.A., Unnasch T.R.;
RT "Epitopes of the Onchocerca volvulus RAL1 antigen, a member of the
RT calreticulin family of proteins, recognized by sera from patients
RT with onchocerciasis."
RL Infect. Immun. 62:3696-3704(1994).
RN [2]
RP SEQUENCE OF 53-388 FROM N.A.
RX MEDLINE=88273584; PubMed=2455736;
RA Unnasch T.R., Gallin M.Y., Soboslay P.T., Erttmann K.D., Greene B.M.;
RT "Isolation and characterization of expression cDNA clones encoding
RT antigens of Onchocerca volvulus infective larvae."
RL J. Clin. Invest. 82:262-269(1988).
CC !- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20565; AAA59056.1; -
DR PIR: A32507; A32507; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Calcium-binding; Repeat; Antigen; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 388 RAL-1 PROTEIN.
FT DOMAIN 189 253 4 X APPROXIMATE REPEATS.
FT REPEAT 189 200 1-1.
FT REPEAT 208 219 1-2.
FT REPEAT 225 236 1-3.
FT REPEAT 242 253 1-4.
FT DOMAIN 257 295 3 X APPROXIMATE REPEATS.
FT REPEAT 257 267 2-1.
FT REPEAT 271 281 2-2.
FT REPEAT 285 295 2-3.
FT DOMAIN 353 388 ARG/LYS-RICH (BASIC).
FT DISULFID 135 161 BY SIMILARITY.
SQ SEQUENCE 388 AA; 45298 MW; 9537F298A2D31CD6 CRC64;

Query Match 61.9%; Score 1366.5; DB 1; Length 388;
Best Local Similarity 64.8%; Pred. No. 9.4e-72; Mismatches 69; Indels 9; Gaps 4;
Matches 241; Conservative 53;

```

QY 4 VYFKEQFLDGGWTSRWIESKHKSDFGKFLVSSGKEYGDEEKDKLQTSQDAREFVALSAS 63
DB 19 IYFKEFSD-DWKEKRWIKSKHKDDFGKWEISHGKFYGDVAVKDKGLKTQDAKFYSIGAK 77
QY 64 FE-PFSNKGQTLVQVFTVKHQBNDCCGGYVKLFPSNLSQDTDMHGDSYNTMFGPDICGP 122
DB 78 FDKSFSNKGSLVIOFSYKHQBODIDCGGGYVKLMASDVNDSHGEPYHMFPGDICGP 137
QY 123 GTKKVVHIFNYKGNVLINKIRCKDDEFTHLYTLVIRPDNTYEVKIDNSQVESGSLEDD 182
DB 138 GTKKVVHIFNYKGNVLINKIRCKDDEFTHLYTLVIRPDNTYEVKIDNSQVESGSLEDD 197
QY 183 WDFLPPKKIKDPDASKPDERAKIDPDTSKPEDWDPKPHIIPDPAKPEDDDEEMDG 242
DB 198 WDFLPPKKIKDPDASKPDERAKIDPDTSKPEDWDPKPHIIPDPAKPEDDDEEMDG 257
QY 243 EWEPPVIONPEYKGEWKPQIDNPDKYKGTWTHPEIDNPESPDPSIYAYDNFVGLDLW 302
DB 258 EWEPPVIONPEYKGEWKPQIDNPDKYKGTWTHPEIDNPESPDPSIYAYDNFVGLDLW 317
QY 303 QVKSCTIFDNFLINDAYAEFEENETWGVTKAAEKQKQKDEORLKEEED----- 356
DB 318 QVKSCTIFDNFLINDAYAEFEENETWGVTKAAEKQKQKDEORLKEEED----- 376
QY 357 KKRKEEAEDEK 368
DB 377 KKRKRANRKKK 388

RESULT 10
CRTC_RICCO STANDARD; PRT; 415 AA.
AC P93508;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97435975; PubMed=9290642;
RA Coughlin S.J., Hastings C., Winfrey R. Jr.;
RT "Cloning and characterization of the calreticulin gene from Ricinus
RL communis L.";
RL Plant Mol. Biol. 34:897-911(1997).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DB EMBL; U74631; AAB71420.1; -
DB EMBL; U74630; AAB71419.1; -
DB Medel; 10452; Ricco; 1166; 10452.
DB InterPro; IPR000886; -
DB InterPro; IPR001580; -
DB Pfam; PF00262; calreticulin; 1.
DB PRINTS; PR00626; CALRETICULIN.
DB PROSITE; PS00014; ER_TARGET; 1.
DB PROSITE; PS00803; CALRETICULIN_1; 1.
DB PROSITE; PS00804; CALRETICULIN_2; 1.
DB PROSITE; PS00805; CALRETICULIN_REPEAT; 2.

```

```

KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 415 CALRETICULIN.
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 412 415 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 415 AA; 47522 MW; DD5452E76CC7F8C CRC64;

Query Match 54.7%; Score 1206.5; DB 1; Length 415;
Best Local Similarity 55.7%; Pred. No. 1.4e-62;
Matches 225; Conservative 56; Mismatches 106; Indels 17; Gaps 7;

QY 4 VYFKEQFLDGGWTSRWIESKHKSDFGKFLVSSGKEYGDEEKDKLQTSQDAREFVAL 60
DB 22 VFEERF--EDGWNWVYKSDMKKDKENTAGWNNTSGKNGD-PNDKGIQTSERYFAI 78
QY 61 SASPEPFSNKGTQTLVQVFTVKHQBNDCCGGYVKLFPSNLSQDTDMHGDSYNTMFGPDIC 120
DB 79 SAEPFESNKKDKTLVQFVSVAHEQKLDCCGGYKLLSSSTQKKFGGDTPIYVNLKYGI 138
QY 121 GPQTKVHVIFNYKGNVLINKIRCKDDEFTHLYTLVIRPDNTYEVKIDNSQVESGSLE 180
DB 139 GYSTKKVHAILNYDNTNHLIKKEVPCTDQLTHVTVLIRPDATYSILIDNVKQTSGLY 198
QY 181 DDWFLPPKKIKDPDASKPDERAKIDPDTSKPEDWDPKPHIIPDPAKPEDDDEE 239
DB 199 TDWLLPPKKIKDPDASKPDERAKIDPDTSKPEDWDPKPHIIPDPAKPEDDDEE 258
QY 240 MDGEWPEVIONPEYKGEWKPQIDNPDKYKGTWTHPEIDNPESPDPSIYAYDNFVGL 299
DB 259 EDGWTAPTIANPEYKGPWKPKKIKNPNYKWKAKPMIDNDFDRDDPIYVNLKYGI 318
QY 300 DLWQVKSCTIFDNFLINDAYAEFEENETWGVTKAAEKQKQKDEORLKEEEDKKR 359
DB 319 ELWQVKSCTIFDNFLINDAYAEFEENETWGVTKAAEKQKQKDEORLKEEEDKKR 375
QY 360 KKEEAEDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
DB 376 PADSDADEDDDDADDTEGEDGESKDAEDSAEDV----HDEL 415

RESULT 11
CRTC_NICPL STANDARD; PRT; 416 AA.
AC Q40401;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
GN CALI.
OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE FROM N.A.
RA Porisjuk N., Borisjuk L., Adler K., Sitaillio L., Tewes A.,
RA Manteuffel R.;
RT "Differential expression of calreticulin during somatic and
RL zygotic embryogenesis of Nicotiana plumbaginifolia.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

entties requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL: Z71395; CAA95999.1; -;
CC Wende; 9149; Nucle; 1166; 9149.
DR InterPro: IPR001580; -;
DR InterPro: IPR001580; -;
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00014; ER.TARGET; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 416 CALRETICULIN.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 413 416 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 416 AA; 47481 MW; 5026F3152B8828C0 CRG64;

Query Match 53.1%; Score 1171.5; DB 1; Length 416;
Best Local Similarity 53.2%; Pred. No. 1.4e-60;
Matches 216; Conservative 59; Mismatches 104; Indels 27; Gaps 7;

QY 4 VYFKEQFLDGDGWTWIESKHKSD---FGKFVLSGKFGVDEKDKGLQTSODARFYAL 60
Db 29 VFEESEF--NDGWESRWKSEWKKDENMAGEWNTSGKNGD-ANDKGQTSEDYRFYAI 85
QY 61 SASFEFSNKGQTLVQVTVKHEQNTDCGGYVYKLPNSLDQTMHGDSEYNTMFGPDIC 120
Db 86 SAEFEFSNKGKLVQFVSVKHEQKLDCCGGYMKLLSGVDQKKFGGDTPTYSIMFGPDIC 145
QY 121 GPCTKKVHVFNKGNVNLINKDIRCKDDEFTHLTVLRPDNTYEVKIDNSOVESGSL 180
Db 146 GYSTKKVHAILNNTNNTNLIKDVPCETDQLTHVYFIRPDATYSLILDNLEKQTSGLY 205
QY 181 DWDVFLPPKKIKDPDASKPDMDERAKIDDPDTSKPEDWDK-PEHIPDPDAKPEDWDE 239
Db 206 SDWDLPLPAKKIKDPEAKKPEDWDEKEFIDDPEDKPEGYDDIPPEITDPDAKPEDWDE 265
QY 240 MGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWVHPEINPEYSPDPSPYAYDNFVGL 299
Db 266 EDGEWAPTIPNPYKGEWKPKKIKNPYKWKAPLIDNPDKDPDPLVYFVKLYGV 325
QY 300 DLQVQKSGTIFDNFLITNDEAYAEFGNETGWYTKAAEQMKDKQDEORLKEEEDKKR 359
Db 326 ELQVQKSGTIFDNVLCDDPEYAKATAETWG-----KQDAKAEAEAEKKR 374
QY 360 KEEEA-----EDKEDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
Db 375 EEEESKAAPADSAEEDDDADDSDDDADDKSESDD----EAHDEL 416

RESULT 12
CRTC_PUAR STANDARD; PRT; 421 AA.
AC Q9XF98;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE CALRETICULIN PRECURSOR.
OS Prunus armeniaca (Apricot).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Rosales; Rosaceae; Prunus.
OX NCBI_TaxID=36596;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BERGERON; TISSUE=Mesocarp, and Endocarp;
RA Mbeque-A-Mbeque D., Fils-Lycan B.R.;
RT "Molecular cloning and nucleotide sequence of a calreticulin from

RT apricot (Prunus armeniaca cv. Bergeron).";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

Query Match 53.0%; Score 1169.5; DB 1; Length 421;
Best Local Similarity 53.8%; Pred. No. 1.9e-60;
Matches 217; Conservative 63; Mismatches 110; Indels 13; Gaps 7;

QY 4 VYFKEQFLDGDGWTWIESKHKSD---FGKFVLSGKFGVDEKDKGLQTSODARFYAL 60
Db 26 VFEESEF--EDGWKRWVTSWKKNLAGEWNTSGKNGD-PNDKGQTSEDYRFYAI 82
QY 61 SASFEFSNKGQTLVQVTVKHEQNTDCGGYVYKLPNSLDQTMHGDSEYNTMFGPDIC 120
Db 83 SAEFEFSNKGKTLVQFVSVKHEQKLDCCGGYMKLLSGVDQKKFGGDTPTYSIMFGPDIC 142
QY 121 GPCTKKVHVFNKGNVNLINKDIRCKDDEFTHLTVLRPDNTYEVKIDNSOVESGSL 180
Db 143 GYSTKKVHAILNNTNNTNLIKDVPCETDQLTHVYFIRPDATYSLILDNLEKQTSGLY 202
QY 181 DWDVFLPPKKIKDPDASKPDMDERAKIDDPDTSKPEDWDK-PEHIPDPDAKPEDWDE 239
Db 203 SDWDLPLPAKKIKDPEAKKPEDWDEQYIPDPEDKPEGYDDIPPEITDPDAKPEDWDE 262
QY 240 MGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWVHPEINPEYSPDPSPYAYDNFVGL 299
Db 263 EDGEWAPTIPNPYKGEWKPKKIKNPYKWKAPLIDNPDKDPDPLVYFVKLYGV 322
QY 300 DLQVQKSGTIFDNFLITNDEAYAEFGNETGWYTKAAEQMKDKQDEORLKEEEDKKR 359
Db 323 ELQVQKSGTIFDNVLCDDPEYAKATAETWGKQDAKAE-----AFEELEKKLEESKED 379
QY 360 KEEEA-----EDKEDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
Db 380 PVDSDAED-ADNEAEDGESDSKSPDSTEESAETAEKHDEL 421

RESULT 13
CRTC_ARATH STANDARD; PRT; 424 AA.
AC Q38858; O80486; O04152;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)


```
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 401 CALRETICULIN.
FT SITE 398 401 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 401 AA; 45910 MW; 056B074C16292674 CRC64;

Query Match 52.8%; Score 1164.5; DB 1; Length 401;
Best Local Similarity 54.1%; Pred. No. 3.5e-60;
Matches 209; Conservative 72; Mismatches 98; Indels 7; Gaps 6;

QY 4 VYKFEQFLDGDGCHTWSWIESKHKSD--FGKFVLSGKFGYGDDEKDKGLQTSQDARFYALSAS 63
DB 20 IYKTEF--EPDWETRWTHSTAKSDYKGLTSGKFGYGDKAKDAGTQTSQDARFYAISP 77
QY 64 F-BPFGSKGOTLVVQFTVKHEQNIDCGGVYKLPSPNSLDQTDHMGDSEYNIMFGPDICGP 122
DB 78 IASFSNECKDLVLFQSVKHEQIDCGGVYKLLP-SVDAAKFTGTPYHIMFGPDICG- 135
QY 123 GTKKVVHVFYKGNVLIINKDCKDEFTHLVTLVRPNTYEVKIDNSQVSGSLED 182
DB 136 ATKKIIFILTYKGNLWKRKEPRCETDLSHTVTAVIKADRTVEVLVDQVKRESGTLEED 195
QY 183 WFLPKKIKDPDASKPEWDERAKIDDDPTDSKPEDWDR-PEHIPDPDAKKPEDWDEMD 241
DB 196 WEILKPKTIPDPEDKKPADWDFDMDVDPEDKKPEDWDRPEAIPDPDATQPDWDEED 255
QY 242 GEWEPPVIONPEYKGEWKPRQIDNPYKGTWVHPEIDNPESYDPSPSIYAYDNFVGLGL 301
DB 256 GKWEAPWISNPKYKGEWAKKIPNPAYKGVKPRDIPNPYEYADDAKHFDELAAGFDL 315
QY 302 WQVKSITFDNFLTNDENAYAEFGNETGWVTKAAEKQMKODEQRLKEEEDKKRKE 361
DB 316 WQVKSITFDNIIVTDSLAEAKAFYQDTNGATKDAEKKAFDSAEADR-KKERDERKKQE 374
QY 362 EEEAEDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 387
DB 375 EEEKTAEEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400

RESULT 15
CRTC.BETVU STANDARD; PRT: 416 AA.
AC 081919;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
OC Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VV-D/RS5; TISSUE-Leaf;
RA Viereck R.;
RT "Nucleotide sequence from sugar beet calreticulin.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ002057; CAA05161.1; -
DR Mendel; 32549; Betvu.1166;32549.
```

```
DR InterPro: IPR000886; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN
DR PROSITE: PS00814; ER_TARGET; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 416 CALRETICULIN.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SITE 413 416 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 416 AA; 48136 MW; 565FEC3489F77CA7 CRC64;

Query Match 52.7%; Score 1161.5; DB 1; Length 416;
Best Local Similarity 53.6%; Pred. No. 5.3e-60;
Matches 215; Conservative 62; Mismatches 109; Indels 15; Gaps 7;

QY 4 VYKFEQFLDGDGCHTWSWIESKHKSD--FGKFVLSGKFGYGDDEKDKGLQTSQDARFYAL 60
DB 27 VFEEERF--EDGWEKRWKSEWKKDESMAGENYTSKGWNGD-ANDKGIQTSYEDRYAI 83
QY 61 SASFEFNSKGTLLVQFTVKHEQNIDCGGVYKLPSPNSLDQTDHMGDSEYNIMFGPDIC 120
DB 84 SAEFPFESNKDNTLVQFSVKHEQKLDCCGGYMKLLSGEVDQKKFGDTPYSIMFGPDIC 143
QY 121 GPGTKKVVHVFYKGNVLIINKDCKDEFTHLVTLVRPNTYEVKIDNSQVSGSLE 180
DB 144 GYSTKKVHAIFNYNDTNHLIKKDVPCETDQLTHVYTFILRPDATYSILIDNQEKQTGSLY 203
QY 181 DWDNLFPPKKIKDPDASKPEWDERAKIDDDPTDSKPEDWDR-PEHIPDPDAKKPEDWDEE 239
DB 204 TWDLPLPARKIKDPDAKKPEDWDRKFIQDPEDKKPEGYDDIPAEITDPEAKKPEDWDEE 263
QY 240 MDGEWPPVIONPEYKGEWKPRQIDNPYKGTWVHPEIDNPESYDPSPSIYAYDNFVGLGL 299
DB 264 EDGEWTAPTIPNPYKGPWKAKKIPNPYKGVKWKAPMIDNPEFKDDPELYVVPKLYRGV 323
QY 300 DLWQVKSITFDNFLTNDENAYAEFGNETGWVTKAAEKQMKODEQRLKEEEDKKR 359
DB 324 ELWQVKSITFDNVLVCDDEPEYAKQLAEETWKGQKDAEKA---AFEELEKKREEEETKDD 380
QY 360 KEEEEAEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
DB 381 PVESDAED-EDAEADDSKDDADSKDDKDDDD-----QHDEL 416
```

Search completed: October 21, 2001, 03:08:24
Job time: 274 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:04:25 ; Search time 92.24 Seconds
(without alignments)
573.743 Million cell updates/sec

Title: US-09-807-148-3
Perfect score: 2206
Sequence: 1 EPAYVFEQFLDGGWTSRW.....EEDKEDEEDVPGQAKDEL 400

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_mammal:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1837.5	83.3	411	13 Q91710	Q91710 xenopus lae
2	1775.5	80.5	419	13 Q98984	Q98984 rana rugosa
3	1687	76.5	405	5 Q26268	Q26268 aplysia cal
4	1662	75.3	417	13 Q9PUC1	Q9PUC1 brachydanio
5	1639	74.3	343	13 Q91711	Q91711 xenopus lae
6	1585.5	71.9	410	5 Q16893	Q16893 amblyomma a
7	1510.5	68.5	421	5 Q9U6S0	Q9U6S0 strongyloce
8	1474	66.8	403	5 Q76961	Q76961 necator ame
9	1456.5	66.0	406	5 Q9U916	Q9U916 drosophila
10	1427.5	64.7	375	5 Q18478	Q18478 litomosoid
11	1397.5	63.3	387	5 Q97372	Q97372 dirofilaria
12	1350	61.2	318	13 Q9PTX7	Q9PTX7 lampetra re
13	1251.5	56.7	321	13 Q9U5G0	Q9U5G0 eptatretus
14	1201	54.4	421	10 Q43712	Q43712 zea mays (m
15	1176.5	53.3	427	10 Q9FYV2	Q9FYV2 pinus taeda
16	1175	53.3	412	10 Q40040	Q40040 hordeum vul
17	1175	53.3	415	10 Q40041	Q40041 hordeum vul
18	1165.5	52.8	389	10 Q40567	Q40567 nicotiana t
19	1103.5	50.0	396	5 Q45034	Q45034 schistosoma

20 1057.5 47.9 350 5 Q26514 schistosoma
21 996 45.1 321 10 Q41799 zea mays (m
22 992 45.0 422 10 Q22502 brassica na
23 979.5 44.4 214 4 Q9UDG2 homo sapien
24 806 36.5 403 5 Q9XYF8 trypanosoma
25 799.5 36.2 401 5 Q9U9N9 trypanosoma
26 796 36.1 622 13 Q98985 rana rugosa
27 747 33.9 240 10 Q9ST29 solanum mel
28 735 33.3 291 5 Q9YIV1 leishmania
29 725 32.9 397 5 Q94592 leishmania
30 673.5 30.5 582 5 Q4702 schistosoma
31 649 29.4 582 5 Q9TFV3 schistosoma
32 647 29.3 582 5 Q76214 schistosoma
33 641 29.1 582 3 Q9HFC6 yarrowia li
34 630 28.6 583 5 Q9VXF6 drosophila
35 626.5 28.4 559 5 Q9NG26 tritrichomo
36 622.5 28.2 605 5 Q02393 drosophila
37 608 27.6 545 5 Q9VAL7 drosophila
38 606 27.5 532 10 Q9LY26 arabidopsis
39 596 27.0 556 5 Q9I7S9 drosophila
40 578.5 26.2 543 5 Q9VVP4 drosophila
41 551 25.0 428 10 Q41798 zea mays (m
42 500.5 22.7 297 5 Q24097 drosophila
43 322 14.6 272 4 Q16094 homo sapien
44 302.5 13.7 101 10 Q40751 parthenium
45 251 11.4 70 10 Q9SAW3 lithospermum

ALIGNMENTS

RESULT 1

Q91710 ID Q91710 PRELIMINARY; PRT; 411 AA.
AC Q91710;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CALRETICULIN PRECURSOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=93074997; Pubmed=1445218;
RA Traves S., Torzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous
system";
RL Biochem. J. 287:579-581(1992).
DR EMBL; X67597; CAA47866.1; -;
DR InterPro; IPR000886; -;
DR InterPro; IPR001580; -;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PS00626; CALRETICULIN.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 12 POTENTIAL.
FT CHAIN 13 411 CALRETICULIN.
SQ SEQUENCE 411 AA; 48344 MW; 891DA66E0EBBEFA CRC64;

Query_Match 83.3%; Score 1837.5; DB 13; Length 411;
Best%Local Similarity 80.5%; Pred. No. 2e-120;
Matches 323; Conservative 44; Mismatches 31; Indels 3; Gaps 2;
QY 1 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFGVLLSSGKFGYDEKDKGLQTSQDARFAL 60

[illegible]

	Qy	304	VKSTIFDNFLITNDEAYAEFGNETWGVTKAAEKOMKKQDKODEQRILKEBEDKKRKEE	363
	Dd	321	VKSITFDNFLLITDDVEAEKFGTDTW-VQRTRTEMKQQEEERKKREEEKSKKDON	379
	Qy	364	EAEDEKEDOK-DEDEDEDEKDEEDVPGAKDEL	400
	Dd	380	EDEDEDEDEPEDDHTTEPPPEEBEGDALPKDEL	417
	RESULT	5		
	Q91711		PRELIMINARY; PRT; 343 AA.	
	ID	A091711;		
	OC	01-JAN-1998 (TrEMBLrel. 05, Created)		
	DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)		
	DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)		
	DE	CALRETICULIN (FRAGMENT).		
	OS	Xenopus laevis (African clawed frog).		
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;		
	OC	Xenopodinae; Xenopus.		
	OX	NCLTaxID=8355;		
	RN	{1}		
	SEQUENCE	FROM N.A.		
	RC	TISSUE=BRAIN;		
	RX	MEDLINE=93074997; PubMed=1445218;		
	RA	Treves S., Zorzato F., Pozzan T.;		
	RT	"Identification of calreticulin isoforms in the central nervous		
	FT	system";		
	RL	Biochem. J. 287:579-581(1992).		
	DR	ENBL; X67598; CAA47867.1; -.		
	DR	InterPro; IPR001580; -.		
	DR	Pfam; PF00262; calreticulin.1.		
	DR	PRINTS; PR00626; CALRETICULIN.		
	DR	PROSITE; PS00803; CALRETICULIN_1; 1.		
	DR	PROSITE; PS00804; CALRETICULIN_2; 1.		
	DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.		
	FT	NON_TER		
	FT	NON_TER		
	SEQUENCE	343 AA; 40105 MW; 3E7DDAF3B91DE1 CRC64;		
	Query Match	74.3%; Score 1639; DB 13; Length 343;		
	Best Local Similarity	83.7%; Pred.No.1.le-106;		
	Matches	283; Conservative 31; Mismatches 24; Indels 0; Gaps		
	Qy	12	DGDGWTSRWTESKHSDPFGKFVLSGGKFYGDGEKDKGLQTSDARFYALSASEPFSNGK	71
	Dd	1	DGDGWTORWESKHSDYGKFKLSAGKFYGDSEKDKGLQTSDARFYAMSSRFESFNKD	60
	Qy	72	QTLVVQFVTVKHEQNDCGGGVKVLFPNSLDQTMHGDSEYNMFGPDICGPGTKKVHVIF	131
	Dd	61	QTLVIQSVKHEQNDCGGGVKLVFPADLEQTEHGESEYNMFGPDICGPTKKVHVIF	120
	Qy	132	NYKGKNVLNKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGLEDDMDFLPKKI	191
	Dd	121	QYKKKNLQINKIRCKDDSFTHLYTLIVRPDNTYEVKIDNSKVESGLEDDMDFLPKKI	180
	Qy	192	KDPDAKPEDWDERAKTDSTDSPEDWDKPEHIPDPDAKKPEDWDEMDEMPPIQN	251
	Dd	181	KDPAKPPDWDERAPKIDDPEDKKPEDWEKPEHIPDPDAVKPEDWDEMDEMPPIQN	240
	Qy	252	PEYGEWKPRQINDPYKGTWIHPIDEINPEYSPPDISIYADNFGVLGLDLWQVKS	311
	Dd	241	PDLOGEWKPRQINDPYKGRWIHPDMNDPEYTDSITLYSYESFGVIGLWLWQVKS	300
	Qy	312	NFLTNDDEAYAEFGNETWGVTKAAEKOMKKQDKODEQR	349
	Dd	301	NFLTNDDEKYAEYGNETWGVTKAAEKMKKEQDEEDR	338

RESULT
Q16893

```

ID Q16893 PRELIMINARY; PRT; 410 AA.
AC Q16893;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CALRETICULIN.
GN CRT-1.
OS Amblyomma americanum.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Amblyomma.
OX NCBI_TaxID=6943;
RN [1]
RP SEQUENCE OF 49-410 FROM N.A.
RC TISSUE-SALIVARY GLANDS;
RA Jaworski D.C., Simmen F.A., Lamoreaux W.J., Coons L.B., Muller M.T.,
RT Needham G.R.;
RT "A secreted calreticulin protein in Ixodid tick (Amblyomma americanum)
RT saliva.";
RL J. Insect Physiol. 41:369-375(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-SALIVARY GLANDS;
RA Jaworski D.C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-SALIVARY GLANDS;
RA Fain-Thornton J.M., Jaworski D.C., Needham G.R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: 007708; AAC79094.1; -.
DR InterPro: IPR000886; -.
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PRODOM: PD001866; -.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 410 AA; 47485 MW; 32CCB8750A17DC54 CRC64;

Query Match 71.9%; Score 1585.5; DB 5; Length 410;
Best Local Similarity 70.7%; Pred. No. 7e-103;
Matches 282; Conservative 38; Mismatches 68; Indels 11; Gaps 2;

QY 1 EPAYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSODARFYAL 60
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 17 DPTVYFKEEFDGSGKSRWVSTKGNLKGKFLVLSAGKFGYDAEKSGLQTSDFYGI 76
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 61 SASFEPSNKGQTLVQFTVKHEQNTDCGGYVKLPNSLDQTDHMGDSYINMFGPDIC 120
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 77 SSKFEPSNEGKTLVQFTVKHEQNTDCGGYVKLPNSLDQTDHMGDSYINMFGPDIC 136
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 121 GPTCKVHVIFNYKGNVLINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSLE 180
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 137 GPTCKVHVIFNYKGNVLINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSLE 196
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 181 DWDVLPKKIKDPDASKPDWDERAKIDDDPTDPSKPDWDERAKIDDDPTDPSKPDWDERAK 240
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 197 SDWSFLPPKKIKDPDASKPDWDERAKIDDDPTDPSKPDWDERAKIDDDPTDPSKPDWDERAK 256
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 241 DGEWEPVQNPEYKGEWKPRQIDNPYKGTWHPIDNPYSPDPSPYAYDNFVGLGLD 300
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 257 DGEWEPVQNPEYKGEWKPRQIDNPYKGTWHPIDNPYSPDPSPYAYDNFVGLGLD 316
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 301 LWQVKSGLTFDNLITNDREAYAEFGNETGWTKAAEKQKMDKQDEQRKKEEEDKKRK 360
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 317 LWQVKSGLTFDNLITNDREAYAEFGNETGWTKAAEKQKMDKQDEQRKKEEEDKKRK 369
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 361 EEEAEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 399
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 370 KEDDAK----DEDEFEDEEKEDEKEDEEETTAAPDEDD 404
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:

```

```

RESULT 7
QYU6S0 PRELIMINARY; PRT; 421 AA.
AC QYU6S0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CALRETICULIN PRECURSOR.
GN CALRET.
OS Strongylocentrotus purpuratus (purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Euechinozoa; Echinacea; Echinacea; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RA Susan J.M., Just M.L., Lennarz W.J.;
RT "Cloning and Characterization of AlphaP Integrin and Calreticulin in
RT Embryos of the Sea Urchin.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF177915; AAD55725.1; -.
DR InterPro: IPR000886; -.
DR InterPro: IPR001580; -.
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 421 CALRETICULIN.
SQ SEQUENCE 421 AA; 48822 MW; 172C664F59F41F93 CRC64;

Query Match 68.5%; Score 1510.5; DB 5; Length 421;
Best Local Similarity 66.5%; Pred. No. 1.2e-97;
Matches 270; Conservative 58; Mismatches 65; Indels 13; Gaps 5;

QY 4 VYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSODARFYALSA 62
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 20 VYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSODARFYALSA 78
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 63 SFEFPSNKGQTLVQFTVKHEQNTDCGGYVKLPNSLDQTDHMGDSYINMFGPDICGP 122
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 79 KTFDFSNECKDLVQFTVKHEQNTDCGGYVKLPNSLDQTDHMGDSYINMFGPDICGP 138
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 123 GTKKVVHVIYKGNVLINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSLED 182
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 139 GTKKVVHVIYKGNVLINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSLED 198
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 183 WDFLPPKKIKDPDASKPDWDERAKIDDDPTDPSKPDWDERAKIDDDPTDPSKPDWDERAK 242
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 199 WDFLPPKKIKDPDASKPDWDERAKIDDDPTDPSKPDWDERAKIDDDPTDPSKPDWDERAK 258
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 243 EWEPPVQNPEYKGEWKPRQIDNPYKGTWHPIDNPYSPDPSPYAYDNFVGLGLDLW 302
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 259 EWEPPVQNPEYKGEWKPRQIDNPYKGTWHPIDNPYSPDPSPYAYDNFVGLGLDLW 318
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 303 QYKSGTIFDNLITNDREAYAEFGNETGWTKAAEKQKMDKQDEQRKKEEEDKKRKEE 362
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 319 QYKSGTIFDNLITNDREAYAEFGNETGWTKAAEKQKMDKQDEQRKKEEEDKKRKEE 378
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 363 --EAEADKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 400
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
QY 379 GGDDEGDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 421
Db : |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:

RESULT 8
QYU6S0 PRELIMINARY; PRT; 403 AA.
AC QYU6S0;
DT 01-NOV-1998 (TREMBLrel. 08, Created)

```


DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS: PS00826; CALRETICULIN.1.
 DR PROSITE: PS00803; CALRETICULIN.1; 1.
 DR PROSITE: PS00804; CALRETICULIN.2; 1.
 DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
 FT NON_TER 375 375
 SQ SEQUENCE 375 AA; 43842 MW; 03F7642F8FF7A5B8 CRC64;

Query Match 64.7%; Score 1427.5; DB 5; Length 375;
 Best Local Similarity 69.3%; Pred. No. 6.2e-92;
 Matches 251; Conservative 46; Mismatches 56; Indels 9; Gaps 3;

QY 4 VYFKEQFLDGDGWTSSWIESKHKSDFKFLVSSGKFGYDEEDKGLQTSODARFYALSAS 63
 Db 19 VYFKEEFLD-DWENRWIKSKHKDDFGKWEISHGKFGYDAVKDGLTKTQDAKFYSIGAK 77
 QY 64 FEP-PSNKQTLVQFTVYKHEONIDCGGKYLFPNSLDQTMHGDSEYNIMFGPDICGP 122
 Db 78 FDKGFSNKGKSLVIOFSVHKEQIDCGGKYLFPNSLDQTMHGDSEYNIMFGPDICGP 137
 QY 123 GTKKVVHVFYKGNVINKDRCRDETHLYTLVVRPNTYEVKIDNSQVSGSLEDD 182
 Db 138 GTKKVVHVFYKGRNMIKKDRCRDETHLYTLVVRPNTYEVKIDNSQVSGSLEDD 197
 QY 183 WFLPPKKTKDPDASKPEDWDERAKIDDPDTSKPEDWDPKPEHIPDPDAKKPEDWDEMDG 242
 Db 198 WFLPPKKTKDPDASKPEDWDEREVIDDDKKPEDWDPKPEHIPDPDAKKPEDWDEMDG 257
 QY 243 EWEPPVIONPEYKGEWKPRQIDNPYKGTWIIHPEIDNPYSPDPSIYAYDNFGVLGLDLW 302
 Db 258 EWEPPVMDNPYKGEWKPKQKNPAYKGIWIIHPEIDNPYSPDPSIYAYDNFGVLGLDLW 317
 QY 303 QVKSGETIFDNFLITNDEAYAEFGNETGVTGVTAAEKOMKODEEORLKEEEDKKRKEE 362
 Db 318 QVKSGETIFDDIIVTSTEEAKFGEKTLKTKQGEKKKKEKQD-----EKEKKRKEE 370

QY 363 EE 364
 Db 371 EE 372

RESULT 11
 O97372
 ID O97372 PRELIMINARY; PRT; 387 AA.
 AC O97372;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE CALRETICULIN PRECURSOR.
 OS Dirofilaria immitis.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Dirofilaria.
 OX NCBI_TaxID=6287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99094497; PubMed=9879888;
 RA Tsuji N., Morales T.H., Ozols V.V., Carmody A.B., Chandrashekar R.;
 RT "Molecular characterization of a calcium-binding protein from the
 filarial parasite Dirofilaria immitis."
 RL Mol. Biochem. Parasitol. 97:69-79(1998).
 DR EMBL: AF052978; AAD03405.1; -;
 DR InterPro: IPR001580; -;
 DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS: PS00826; CALRETICULIN.
 DR PRODOM: PD001866; -; 1.
 DR PROSITE: PS00803; CALRETICULIN.1; 1.
 DR PROSITE: PS00804; CALRETICULIN.2; 1.
 DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
 KW SIGNAL.
 FT CHAIN 1 18 POTENTIAL.
 FT CHAIN 19 387 CALRETICULIN.
 SQ SEQUENCE 387 AA; 44941 MW; E7741BF6AFA5885 CRC64;

Query Match 63.3%; Score 1397.5; DB 5; Length 387;
 Best Local Similarity 65.6%; Pred. No. 7.9e-90;
 Matches 244; Conservative 55; Mismatches 64; Indels 9; Gaps 3;

QY 4 VYFKEQFLDGDGWTSSWIESKHKSDFKFLVSSGKFGYDEEDKGLQTSODARFYALSAS 63
 Db 19 VYFKEEFS-DWENRWIKSKHKDDFGKWEISHGKFGYDAVKDGLTKTQDAKFYSIGAK 77
 QY 64 FEP-PSNKQTLVQFTVYKHEONIDCGGKYLFPNSLDQTMHGDSEYNIMFGPDICGP 122
 Db 78 FDKGFSNKGKSLVIOFSVHKEQIDCGGKYLFPNSLDQTMHGDSEYNIMFGPDICGP 137
 QY 123 GTKKVVHVFYKGNVINKDRCRDETHLYTLVVRPNTYEVKIDNSQVSGSLEDD 182
 Db 138 GTKKVVHVFYKGRNMIKKDRCRDETHLYTLVVRPNTYEVKIDNSQVSGSLEDD 197
 QY 183 WFLPPKKTKDPDASKPEDWDERAKIDDPDTSKPEDWDPKPEHIPDPDAKKPEDWDEMDG 242
 Db 198 WFLPPKKTKDPDASKPEDWDEREVIDDDKKPEDWDPKPEHIPDPDAKKPEDWDEMDG 257
 QY 243 EWEPPVIONPEYKGEWKPRQIDNPYKGTWIIHPEIDNPYSPDPSIYAYDNFGVLGLDLW 302
 Db 258 EWEPPVMDNPYKGEWKPKQKNPAYKGIWIIHPEIDNPYSPDPSIYAYDNFGVLGLDLW 317
 QY 303 QVKSGETIFDNFLITNDEAYAEFGNETGVTGVTAAEKOMKODEEORLKEEEDKKRKEE 362
 Db 318 QVKSGETIFDDIIVTSTEEAKFGEKTLKTKQGEKKKKEKQD-----EKEKKRKEE 370

QY 363 EBAEDKEDDEK 374
 Db 371 KRPPKKKKKKKK 382

RESULT 12
 O9PTX7.
 ID O9PTX7 PRELIMINARY; PRT; 318 AA.
 AC O9PTX7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE CALRETICULIN (FRAGMENT).
 OS Lampetra reissneri.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Lampetra.
 OX NCBI_TaxID=7753;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20063780; PubMed=10594174;
 RA Kuraku S., Hoshiyama D., Katoh K., Suga H., Miyata T.;
 RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded
 genes".
 RL J. Mol. Evol. 49:729-735(1999).
 DR EMBL: AB025328; BAA88481.1; -;
 DR InterPro: IPR000886; -;
 DR InterPro: IPR001580; -;
 DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS: PS00626; CALRETICULIN.
 DR PRODOM: PD001866; -; 1.
 DR PROSITE: PS00804; CALRETICULIN.2; 1.
 DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 318 AA; 36997 MW; C88102EALCAC1506 CRC64;

Query Match 61.2%; Score 1350; DB 13; Length 318;
 Best Local Similarity 73.4%; Pred. No. 1.3e-86;
 Matches 234; Conservative 41; Mismatches 40; Indels 4; Gaps 2;

QY 85 NIDCGGKYLFPNSLDQTMHGDSEYNIMFGPDICGPGTKKVVHVFYKGNVINKD 144
 |||||:||||| :|| :|||||: |||||:|||||:|||||:|||||:|||||:|||||

Db 1 NIDCGGYIKLFPKDMQNEBMHGSQYFIMFGPDICGPGTKKVHVIENYKGNHINKDI 60
QY 145 RCRDDETHLYTLIVRPDNTYEVKIDNSQVSGSLEDWDFLPPKKIKDPDAPKPDWDE 204
Db 61 RCRDDETHLYTLIVRPDNTYEVKIDNSQVSGSLEDWDFLPPKKIKDPDAPKPDWDE 120
QY 205 RAKIDDPDTSKPDWKPDPDPAKPPDDEEMGEWEPVQIONPYKGEWKPRQID 264
Db 121 QAKIDDPDTSKPDWKPDPDPAKPPDDEEMGEWEPVQIONPYKGEWKPRQID 180
QY 265 NPYKGTWIIHPEIDNPEYSPDPSIYAYDNFVGLDLQWQVSGTIFDNFLITNDAYAE 324
Db 181 NPYKGTWIIHPEIDNPEYSPDPSIYAYDNFVGLDLQWQVSGTIFDNFLITNDAYAE 240
QY 325 FGNETWGTAAEKQKDKQDEQRLEKEEEDKKRKEEAE---DKEDDEDKDEDEDE 381
Db 241 IGEETWGTAAEKQKDKQDEQRLEKEEEDKKRKEEAE---DKEDDEDKDEDEDEDE 300
QY 382 EDKEEEDVPGQAKDEL 400
Db 301 DEPEDDDDDDP-CLKDEL 318

RESULT 13
QY5G0 PRELIMINARY; PRT; 321 AA.
AC QY5G0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Eptaretus burgeri (Inshore hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniformes;
OC Myxiniidae; Eptaretinae; Eptaretus.
OX NCBI_TaxID=7764;
RN SEQUENCE FROM N.A.
RP TISSUE=LIVER;
RC MEDLINE=20063780; PubMed=10594174;
RX Kuraku S., Hoshiyama D., Katoh K., Suga H., Miyata T.;
RA "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded
RT genes.";
RL J. Mol. Evol. 49:729-735(1999).
DR ENBL; AB025323; BAA88476.1; -;
DR InterPro; IPR000886; -;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866; -; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 1
SQ SEQUENCE 321 AA; 37367 MW; 6E8DFA98D42F7AEF CRC64;

Query Match 56.7%; Score 1251.5; DB 13; Length 321;
Best Local Similarity 68.6%; Pred. No. 9.3e-80;
Matches 221; Conservative 43; Mismatches 49; Indels 9; Gaps 3;
QY 86 IDCGGGYKVLFPNSLDQTMHGSSEYNIMFGPDICGPGTKKVHVIENYKGNHINKDIR 145
Db 2 IDCGGGYKVLFPNSLDQTMHGSSEYNIMFGPDICGPGTKKVHVIENYKGNHINKDIR 61
QY 146 CKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLEDWDFLPPKKIKDPDAPKPDWDE 205
Db 62 CKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLEDWDFLPPKKIKDPDAPKPDWDE 121
QY 206 AKIDDPDTSKPDWKPDPDPAKPPDDEEMGEWEPVQIONPYKGEWKPRQID 265
Db 122 AQIEDTDAKPDWKPDPDPAKPPDDEEMGEWEPVQIONPYKGEWKPRQID 181
QY 266 PDYKGTWIIHPEIDNPEYSPDPSIYAYDNFVGLDLQWQVSGTIFDNFLITNDAYAE 325

Db -182 PKYKGMQHOPEIDNPEYSPDPSIYAYDNFVGLDLQWQVSGSLEDWDFLPPKKIKDPDAPKPDWDE 241
QY 326 GNETWGTAAEKQKDKQDEQRLEKEEEDKKR-----KEEBAEKDEDEDEDEDE 379
Db 242 GROTGWGTAAEKQKDKQDEQRLEKEEEDKKR-----KEEBAEKDEDEDEDEDEDE 301
QY 380 DEEDKE-DEEEDVPGQAKDEL 400
Db 302 EEDKAGESPVEETP--TKDEL 321
RESULT 14
QY43712 PRELIMINARY; PRT; 421 AA.
AC QY43712;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CALCIUM-BINDING PROTEIN PRECURSOR.
GN CRT1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN SEQUENCE FROM N.A.
RP STRAIN=VAR MERIT; TISSUE=ROOT TIP;
RA Napier R.M., Trueman S., Henderson J., Boyce J.M., Hawes C.R.,
ERicker M.D., Venis M.A.;
RL J. Exp. Bot. 46:1603-1613(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96309381; PubMed=8704156;
RA Dresselhaus T., Hagel C., Loerz H., Kranz E.;
RT "Isolation of a full-length cDNA encoding calreticulin from a PCR
RT library of in vitro zygotes of maize.";
RL Plant Mol. Biol. 31:23-34(1996).
DR ENBL; X86772; CAA86728.1; -;
DR EMBL; X89813; CAA61939.1; -;
DR Mendel; 11228; Zeama; 1166; 11228.
DR InterPro; IPR000886; -;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866; -; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Signal; Calcium-binding.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 421 CALRETICULIN.
SQ SEQUENCE 421 AA; 48012 MW; 5AAE02B77ED3126D CRC64;

Query Match 54.4%; Score 1201; DB 10; Length 421;
Best Local Similarity 54.8%; Pred. No. 4.2e-76;
Matches 223; Conservative 67; Mismatches 95; Indels 22; Gaps 8;
QY 4 VYFQEQFLDGGWTSRWIESKHSKSD---FGKFLVSSGKFGYDEEKDKGLQTSQDARFAL 60
Db 27 VYFQEQFLDGGWTSRWIESKHSKSD---FGKFLVSSGKFGYDEEKDKGLQTSQDARFAL 83
QY 61 SASPEFSNKGQTLVQPTVTKHEONIDCGGYKVLFPNSLDQTMHGSSEYNIMFGPDIC 120
Db 34 SAETPEFSNKGQTLVQPTVTKHEONIDCGGYKVLFPNSLDQTMHGSSEYNIMFGPDIC 143
QY 121 GPGKPKVHVIFNYKGNVINKNIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 180
Db 144 GYSTKPKVHVIFNYKGNVINKNIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 203

Search completed: October 21, 2001, 03:10:07
Job time: 342 sec

```
QY 181 DDWDLPPKKIKDPAKPEDWDERAKIDDPDTSKPEDWK-PEHIPDPDAKKPEDWDEE 239
Db 204 EHWDLPPKKIKDPAKPEDWDDKEYIPDPEDKKPEGYDDIPKEIPDPDAKKPEDWDD 263
QY 240 MGEWEPPVIONPEYKGEWKPRQIDNPDKYGTWIIHPEIDNPESDPSPSYAYDNFGLGL 299
Db 264 EGEWTAPTIPPEYKGPWKPRQIKRNPOTGKWKAPMIDNPKDDPIYAFDSLKYIGI 323
QY 300 DLWQVKSGLTIFDNFLITNDAYAAEEFNETGWVTKAAEKQMKDKQDEQRLKEEEDKKR 359
Db 324 ELWQVKSGLTIFDNFLITNDPALAKTFAETGTWGHKEAKAFADEAK-----KKEEDAAK 379
QY 360 ---KEEERAEDEDDKDEDEED---EEDKEEDEDVPGQAKDEL 400
Db 380 GGDEDDLEDEDEDEKADKADSDAEDSKDDEKQ-----HDEL 421
```

RESULT 15

```
Q9FYV2 PRELIMINARY; PRT; 427 AA.
AC Q9FYV2;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE CALRETICULIN.
OS Pinus taeda (Loblolly pine).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OX Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Shealy M.J., Frankis R.C. Jr.;
RT "Complete nucleotide sequence of a cDNA encoding calreticulin from
RL Pinus taeda";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF283816; AAG01147.1; -.
SQ SEQUENCE 427 AA; 49627 MW; 9CD5ED75EEE74F8E CRC64;
```

```
Query Match 53.3%; Score 1176.5; DB 10; Length 427;
Best Local Similarity 54.2%; Pred. No. 2.2e-74;
Matches 224; Conservative 62; Mismatches 102; Indels 25; Gaps 10;

QY 4 VYFKEQFLDGGWTSRWIESKHSD---FGKVLSSGKFYGDDEKDKGLQTSQDARFYAL 60
Db 24 VFEERF---DSSWESRWQSDWKKDESAGDWVHTSGKNGD-PNDKGIQTHTDYFFAI 80
QY 61 SASPEPFSNKGOTLVQFTVKHEQNTDCGGYVVKLFPNSLDQTMHGDSEYNMFQPDIC 120
Db 81 SAAYPEFSNKGOTLVQFSVKEHQKIDCGGYVVKLLSGEIDQKFNFSGETPYSIMFGPDIC 140
QY 121 GPGTKKVHVIFNKGKVNLIKNDIRCKODEFTHLYTLIVRPDNTVEYKIDNSQVESGSL 180
Db 141 GYSTKKVHTILSYKGNP-IKKDVPCEITQDLTHVYTFILRPDNTYSILIDNTDKSGSGLY 199
QY 181 DDWDLPPKKIKDPAKPEDWDERAKIDDPDTSKPEDWK-PEHIPDPDAKKPEDWDEE 239
Db 200 KDMDLLPPKTIKDPNAKPEDWDDKEYIPDPEDKKPEGYDDIPKEIPDPDAKKPEDWDE 259
QY 240 MGEWEPPVIONPEYKGEWKPRQIDNPDKYGTWIIHPEIDNPESDPSPSYAYDNFGLGL 299
Db 260 EGEWTAPTIANPEYKGPWKPRQIKRNPYKWKAPMIDNPKDDPELYVFPNLKYLGI 319
QY 300 DLWQVKSGLTIFDNFLITNDAYAAEEFNETGWVTKAAEKQMKDKQDEQRLKEEEDKKR 359
Db 320 ELWQVKSGLTIFDNFLITNDPEYAKKLAETWAKHKDAE---KEAFDEAKKKEEKEKE 376
QY 360 KEE-----EADKDEDEKDE--DEDEDEKDEE-----EEDVPGQAKDEL 400
Db 377 SDEEDTDEKEKSDDEDADELDDHEKADKKEHDELSEHKEED--KREHDEL 427
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: Octobhr 21, 2001, 03:06:40 ; Search time 42.3 Seconds
(without alignments)
87.618 Million cell updates/sec

Title: US-09-807-148-4
Perfect score: 971
Sequence: 1 EPAAVFEQFLDGDGWTSRW.....PDNTYEVKIDNSQVSGSL 180

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	527	54.3	415	4	US-08-675-816-2
2	290	29.9	593	1	US-08-296-362-2
3	227	5	542	4	US-08-675-816-5
4	89	9.2	660	4	US-09-181-706-8
5	89	9.2	660	4	US-09-458-791-8
6	89	9.2	660	4	US-09-459-066-8
7	77.5	8.0	642	2	US-08-245-511-48
8	77.5	8.0	642	2	US-08-600-993A-48
9	77	7.9	15	3	US-08-946-026-53
10	75	7.7	591	3	US-08-991-408-4
11	75	7.7	1013	2	US-08-866-650-5
12	75	7.7	1013	2	US-09-021-287-5
13	75	7.7	1013	3	US-08-991-408-2
14	72.5	7.5	1484	2	US-08-231-193A-56
15	72.5	7.5	1484	2	US-08-486-273A-56
16	72.5	7.5	1484	3	US-08-940-086A-56
17	72	7.4	263	1	US-08-152-922A-5
18	72	7.4	511	1	US-08-480-604A-20
19	72	7.4	511	2	US-08-405-496A-20
20	72	7.4	608	1	US-08-480-604A-21
21	72	7.4	608	2	US-08-405-496A-21
22	72	7.4	609	1	US-08-480-604A-30
23	72	7.4	984	1	US-08-257-073-3
24	72	7.4	984	2	US-08-184-009-120
25	72	7.4	984	2	US-08-458-356-120
26	72	7.4	2366	1	US-08-480-604A-10
27	72	7.4	2366	2	US-08-405-496A-10

28	71.5	7.4	324	3	US-08-946-914-11	Sequence 11, Appl
29	71.5	7.4	999	2	US-08-473-553A-5	Sequence 5, Appl
30	70	7.2	643	2	US-08-245-511-47	Sequence 47, Appl
31	70	7.2	643	2	US-08-600-993A-47	Sequence 47, Appl
32	70	7.2	992	1	US-07-813-593-2	Sequence 2, Appl
33	70	7.2	992	1	US-07-977-451-2	Sequence 2, Appl
34	70	7.2	992	1	US-07-946-507-2	Sequence 2, Appl
35	70	7.2	992	1	US-08-252-517-2	Sequence 2, Appl
36	70	7.2	992	1	US-07-906-397A-2	Sequence 2, Appl
37	70	7.2	992	1	US-08-601-891-2	Sequence 2, Appl
38	70	7.2	992	2	US-09-021-324-2	Sequence 2, Appl
39	70	7.2	992	5	PCT-US92-02750-2	Sequence 2, Appl
40	70	7.2	992	5	PCT-US92-05401-2	Sequence 2, Appl
41	70	7.2	992	5	PCT-US92-09893-2	Sequence 2, Appl
42	70	7.2	1000	1	US-08-222-299-2	Sequence 2, Appl
43	70	7.2	1000	2	US-08-434-878-2	Sequence 2, Appl
44	70	7.2	1000	5	PCT-US95-03718-2	Sequence 2, Appl
45	70	7.2	3052	2	US-08-557-122A-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-675-816-2
; Sequence 2, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND US
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-575-816-2

Query Match 54.3%; Score 527; DB 4; Length 415;
Best Local Similarity 55.3%; Pred. No. 1.8e-51;
Matches 99; Conservative 31; Mismatches 43; Indels 6; Gaps 3;
QY 4 VYFKEQFLDGDGWTSRWIESKHKSD---FGKVLSSSGKFGYGDKEKGLQTSQDARFYAL 60
Db 22 VFFERF--EDGWNWVKSDWKDKENTAGWNTSGKNGD-PNDKGIQTSYDFRYAI 78
QY 61 SASPEPFSNKGQTLVQVTFVKHEQNICGGYVKLFNPSLQDTDMHGDSSEYNIMFGPDIC 120

```

Db 79 SAEPFESNKKDTLVFQSVKHEQKLDGCGGYMKLLSSSTDQKKFGGDTTPYSIMFGPDIC 138
QY 121 GPCTKVVHVFYKGNKVLNIDKIRCKDDFTHTLYLIVRPDNTYEVKIDNSQVSGSL 179
Db 139 GYSTKKVHAILNVDNHLIKKEVPCETDQTHVYLVIRPDATYSLIDNVEKQTGSL 197

RESULT 2
US-08-296-362-2
; Sequence 2, Application US/08296362
; Patent No. 5691306
; GENERAL INFORMATION:
; APPLICANT: Bergeron, John J.M.
; APPLICANT: Thomas, David Y.
; APPLICANT: Wada, Ikuro
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; TITLE OF INVENTION: PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
; TITLE OF INVENTION: PROTEIN PRODUCTION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/296,362
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deehr, Manya S.
; REGISTRATION NUMBER: 37,120
; REFERENCE/DOCKET NUMBER: 690066.401C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-296-362-2

Query Match 29.9%; Score 290; DB 1; Length 593;
Best Local Similarity 38.1%; Pred. No. 2e-24; Mismatches 65; Indels 34; Gaps 9;
Matches 77; Conservative 26;

QY 4 VYFQELDGDGWTSRWIESKHKSD-----FGKFLVSSGKFGYGDREK-----DKGLQTS 52
Db 70 VYFADF--DRGLTSGWLLSKAKKDDTDDEIAKY---DGKWEVDENKTKLPDGKGLVLM 124
QY 53 QDARFVALSASF--EPFSNKGQTLVQFTVKHEQNIIDCGGYVKKLPFN-----SLDQDTMHG 107
Db 125 SRAKHHAIASAKLNKPELFDTPKPLIVQYVNFQNGIECGGAYVKKLSKTPPELNLDQ--FHD 182
QY 108 DSYNTMFGPDIGCGTKKVVHVFYKGNKVLNIDKIRCK-----DDEFTHTLYL 157
Db 183 KPTTLMFGPDKGE--DYKLHFTFRHNKPTGYVEEKHAKRPADLKTFTDKKHTLYL 241
QY 158 IVRPDNTYEVKIDNSQVSGSL 179
Db 242 ILNPDNSFEILVDQSVNLSGNL 263

```

```

RESULT 3
US-08-675-816-6
; Sequence 6, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND US
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)-622-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-816-6

Query Match 23.4%; Score 227.5; DB 4; Length 542;
Best Local Similarity 36.0%; Pred. No. 2e-17;
Matches 64; Conservative 24; Mismatches 73; Indels 17; Gaps 10;

QY 6 FKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFVALSASF- 64
Db 35 FYESF--EDSFEGRWTLA-KDDY-KGEWKHAKSEGHDE--YGLLVSEKARKYAIYKELD 88
QY 65 EPFSNKGQTLVQFTVKHEQNIIDCGGYVKK-LFPNSILDOT--DMHGDSEYNIMFGPDIC 121
Db 89 EPATLKDGTVLVQFTFRFQNGLECGGAYLKLPQEGAGWTPKDFDNDSPYSIMFGPDKCG 148
QY 122 PGTKKVVHVFYKGNKVLNIDKIRCK-----CKDDEFTHTLYLIVRPDNTYEVKIDNSQ 173
Db 149 -ATNKVHFILKHNKPKSGEYIEHLKYPSPVSKLTHVYVTAALKPDNELRLVDGEE 205

RESULT 4
US-09-181-706-8
; Sequence 8, Application US/09181706
; Patent No. 6130068
; GENERAL INFORMATION:
; APPLICANT: Melanie K. Spriggs, Michael R. Comeau,
; APPLICANT: Robert F. DuBoise, Richard S. Johnson
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle

```


;; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/459,066
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,598
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-459-066-8

Query Match 9.2%; Score 89; DB 4; Length 660;
Best Local Similarity 20.6%; Pred. No. 0.13;
Matches 41; Conservative 25; Mismatches 91; Indels 42; Gaps 6;

QY 5 YFKEQFLDGGWTSRWIESKHKSDFGKVLVLSGKFGDEKDKGLQTSQDAREYALS 61
DB 448 YIAQMLNDEGGFSS-LSSHRWSTFLKVELECDIDGRSVRIIHSRAIKTNDITLYVFF 506
QY 62 ASPEFNSKNGQTLVQFTVKHEQNDICGGYVKLFNLSLDQTDHMGDSEYNIMFGPDICG 121
DB 507 DS--PYSKSALCTYSMAIKHSFSTSKLGCTKQLPSP-----APGICL 548
QY 122 PGTKKV-HVFN-----YKGNVLNKDIRCKDDDEFTHLYTLIVRP 161
DB 549 PAGKVVPHFTFDIEQYNELDDIKPLSQPIFEGSGVKWFIDIKENEHREHYRIYFIKE 608
QY 162 DNTYEVKIDNSQVESGSL 180
DB 609 NTIYSEDTKSKQTRSAQVD 627

RESULT 7
US-08-245-511-48
; Sequence 48, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541

;; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; IMMEDIATE SOURCE:
; CLONE: amia
; FEATURE:
; OTHER INFORMATION: NOTE: the reference contains a
; OTHER INFORMATION: sequence error; the correct sequence shown below is obtain
; OTHER INFORMATION: from GENBANK
; PUBLICATION INFORMATION:
; AUTHORS: Alloing, et al.
; JOURNAL: Mol. Microbiol.
; VOLUME: 4
; PAGES: 633-644
; DATE: 1990
; US-08-245-511-48

Query Match 8.0%; Score 77.5; DB 2; Length 642;
Best Local Similarity 21.2%; Pred. No. 2.4;
Matches 42; Conservative 20; Mismatches 81; Indels 55; Gaps 8;

QY 16 WTSR-----W-----IESKHKSDFGKVLVLSGKFGDEKDKGLQTSQDAREYALSASF 64
DB 164 WNSKLTYSIFWPLNEEFETSKGSDFAKPTDPTSLLYNGFLKGLTAKSSVEF-----VKN 219
QY 65 EPFSNKGQTLVQFTVKH-----EQNDICGG-GYVKLFNLSLDQTDHMGDSEY 111
DB 220 EQYWDKENVHLDITINLAYDGDQESLERNFTSGAYSARLYPTSSNYSKVAEYKDNIIY 279
QY 112 NIMFGPDICGCTKVKHVIFNYKG-----KNVLNKDIR----- 145
DB 280 YTQSGSGIAGLVNIDRQSYNTSKTTDSEKVAATKALLNKDFRQALNFALDRSAYSQI 339
QY 146 -CKDDEFTHLYTLIVRPD 162
DB 340 NGKDGAAVARNLFLVXPD 357

RESULT 8
US-08-600-993A-48
; Sequence 48, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J.
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; IMMEDIATE SOURCE:
; CLONE: amia
; FEATURE:
; OTHER INFORMATION: the reference contains a sequence error; the
; correct sequence shown below is obtained from GENBANK
; PUBLICATION INFORMATION:
; AUTHORS: Alloing, et al.
; JOURNAL: Mol. Microbiol.
; VOLUME: 4
; PAGES: 633-644
; DATE: 1990
; US-08-600-993A-48

Query Match 8.0%; Score 77.5; DB 2; Length 642;
Best Local Similarity 21.2%; Pred. No. 2.4; Mismatches 20; Indels 55; Gaps 8;
Matches 42; Conservative 20;

Qy 16 WTSR-----W-----IESKHKDFGRFVLSSGRFYGDDEKDKGLQTSQARFYALSASF 64
Db 164 WNSKLTYSIFWPLNEEETSGSDFAKPTDPTSLLYNGPFLKGLTAKSSVEF-----VKN 219

Qy 65 EPPFNKSGOTLVVQTVKH-----EQNIDCGG-GYVKLFPPNSLDQT---DMHGDSEY 111
Db 220 EQYWDKENVHLDITNLAYDGSQDESLEARNFTSGAYSVARLYPTSSNYKVAEYKDNII 279

Qy 112 NIMFGPDICGGTKKHVIFNYK-----KNVINKDIR----- 145
Db 280 YTQSGSGIAGLVGNIDRQSNVYKTTTDSKVATKALLNKNDFRQALNFDLRSAYSQAI 339

Qy 146 -CKDDEFTHTLVLRPD 162
Db 340 NGKDGAAVLAVRNLFVKPD 357

RESULT 9
US-08-946-026-53
; Sequence 53, Application US/08946026
; Patent No. 6034218

```

```

; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Mitcham, Jennifer L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,026
; FILING DATE: 07-OCT-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.424C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-946-026-53

Query Match 7.9%; Score 77; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PAVYFKEQFLDGDG 15
Db 2 PAVYFKEQFLDGDG 15

RESULT 10
US-08-946-026-53
; Sequence 4, Application US/08991408
; Patent No. 6008017
; GENERAL INFORMATION:
; APPLICANT: ARLETH, ANTHONY J.
; APPLICANT: WILLETTTE, ROBERT N.
; APPLICANT: ELSHOURBAGY, NABIL A.
; APPLICANT: LI, XIAOTONG
; TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLOID-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

```

;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/991,408
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/034,471
;; FILING DATE: 02-JAN-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: PRESTIA, PAUL F
;; REGISTRATION NUMBER: 23,031
;; REFERENCE/DOCKET NUMBER: ATG-50038
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-407-0700
;; TELEFAX: 610-407-0701
;; TELEX: 846169
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 591 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-991-408-4

Query Match 7.7%; Score 75; DB 3; Length 591;
Best Local Similarity 31.4%; Pred. No. 4.1;
Matches 33; Conservative 15; Mismatches 39; Indels 18; Gaps 7;

QY 68 SNKGOTLVVQF-TVKHEQNIDCGGVYVKLFPPNSLDQTMHGDSEYNIMFGPDICGPG-TK 125
DB 501 SERGSRLELSFQTFEEEDACGYDYVELF-DGLDSTAV-GLGRF-----CGSGPPE 550
QY 126 KHVHFIENYKGNVLI--NKDIRCKDDEFTHLVLRPNTYEVK 168
DB 551 EIVSI-----GDSVLIHFHTDDTINKKGFIHYKSIKIRYPTDTHTKK 591

RESULT 11
US-08-866-650-5
Sequence 5, Application US/08866650
Patent No. 5939321
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,650
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296.93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

;;
;; LENGTH: 1013 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-866-650-5

Query Match 7.7%; Score 75; DB 2; Length 1013;
Best Local Similarity 31.4%; Pred. No. 9;
Matches 33; Conservative 15; Mismatches 39; Indels 18; Gaps 7;

QY 68 SNKGOTLVVQF-TVKHEQNIDCGGVYVKLFPPNSLDQTMHGDSEYNIMFGPDICGPG-TK 125
DB 923 SERGSRLELSFQTFEEEDACGYDYVELF-DGLDSTAV-GLGRF-----CGSGPPE 972
QY 126 KHVHFIENYKGNVLI--NKDIRCKDDEFTHLVLRPNTYEVK 168
DB 973 EIVSI-----GDSVLIHFHTDDTINKKGFIHYKSIKIRYPTDTHTKK 1013

RESULT 12
US-09-021-287-5
Sequence 5, Application US/09021287
Patent No. 5981717
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,287
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/866,650
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296.93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-287-5

Query Match 7.7%; Score 75; DB 2; Length 1013;
Best Local Similarity 31.4%; Pred. No. 9;
Matches 33; Conservative 15; Mismatches 39; Indels 18; Gaps 7;

QY 68 SNKGOTLVVQF-TVKHEQNIDCGGVYVKLFPPNSLDQTMHGDSEYNIMFGPDICGPG-TK 125
DB 923 SERGSRLELSFQTFEEEDACGYDYVELF-DGLDSTAV-GLGRF-----CGSGPPE 972


```

; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,273A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,193
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383B
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1484 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-273A-56

```

```

Query Match 7.5%; Score 72.5; DB 2; Length 1484;
Best Local Similarity 27.7%; Pred. No. 30;
Matches 23; Conservative 13; Mismatches 28; Indels 19; Gaps 4;

QY 45 KDKG--LQTSQARFVALSASPEFSNKGQTLVVQFTVKHEQNIDCGGYV---KLFPS 99
Db 1325 KDKGRFMDGSPYAHMFEMSGESTFANNKSSVP---TAGHHHHNNPGGGYMLSKSLYPD 1381

QY 100 L-----DQTDMDHGDSEY 111
Db 1382 VTQNPFIPTFGDDQCLLGSKSY 1404

```

Search completed: October 21, 2001, 03:06:41
Job time: 331 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:08:24 ; Search time 28.98 seconds
(without alignments)
212.767 Million cell updates/sec

Title: US-09-807-148-4

Perfect score: 971

Sequence: 1 EPAVYKQFLDGDGWSRW.....PDNTYEVKIDNSQVSGSLE 180

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	971	100.0	417	1	CRTC_HUMAN
2	933	96.1	416	1	P27797 homo sapien
3	931	95.9	418	1	P18418 rattus norv
4	930	95.8	400	1	P15253 oryctolagus
5	930	95.8	416	1	P52193 bos taurus
6	699	72.0	406	1	P14211 mus musculus
7	664	68.4	388	1	P29413 drosophila
8	619.5	63.8	395	1	P11012 onchocerca
9	537	55.3	424	1	P27798 caenorhabdi
10	535.5	55.1	401	1	Q23858 dictyosteli
11	531	54.7	416	1	Q82933 euglena gra
12	527	54.3	415	1	O81919 beta vulgar
13	527	54.3	421	1	P93508 ricinus com
14	527	54.3	424	1	Q9xf98 prunus arme
15	522.5	53.8	421	1	Q9slv8 oryza sativ
16	522	53.8	416	1	P42918 bos taurus
17	518	53.3	424	1	Q40401 nicotiana p
18	512	52.7	420	1	Q38858 arabidopsis
19	512	52.7	425	1	Q9std3 chlamydomon
20	509	52.4	416	1	O04151 arabidopsis
21	504	51.9	393	1	Q9zpp1 barberis st
22	494	50.9	420	1	O06814 schistosoma
23	455.5	46.9	424	1	Q9sp22 zea mays (m
24	447	46.0	105	1	O04153 arabidopsis
25	293	30.2	591	1	P28491 sus scrofa
26	290	29.9	591	1	P35564 mus musculu
27	290	29.9	593	1	P35565 rattus norv
28	288	29.7	592	1	P24643 canis famil
29	270	27.8	611	1	P27824 homo sapien
30	261	26.9	610	1	P52194 mus musculu
31	253.5	26.1	560	1	O14967 homo sapien
32	244	25.1	619	1	P36581 schizosacch
33	238.5	24.6	530	1	P34652 caenorhabdi
					P29402 arabidopsis

34	232	23.9	546	1	CALX_SOYBN	Q39817 glycine max
35	215.5	22.2	540	1	CALX_HELTU	Q39994 helianthus
36	215.5	22.2	551	1	CALX_PEA	O82709 pisum sativ
37	202.5	20.9	528	1	CAX2_ARATH	Q38798 arabidopsis
38	133.5	13.7	502	1	CALX_YEAST	P27825 saccharomyc
39	86	8.9	24	1	CRTC_CANFA	P28490 canis famil
40	80.5	8.3	540	1	MTAL_ACICA	P35201 acinetobact
41	79.5	8.2	510	1	PROL_LISMO	P32224 listeria mo
42	78.5	8.1	344	1	ABIC_LACIA	Q01457 lactococcus
43	78	8.0	928	1	HXA2_HAEIN	P45354 haemophilus
44	77.5	8.0	659	1	AMIA_STRPN	P18791 streptococc
45	77	7.9	3130	1	DPOZ_HUMAN	O60673 homo sapien

ALIGNMENTS

RESULT 1					
CRTC_HUMAN					
ID	CRTC_HUMAN	STANDARD:	PRT:	417 AA.	
AC	P27797;				
DT	01-AUG-1992 (Rel. 23, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (52 KDA				
DE	DE RIBONUCLEOPROTEIN AUTOANTIGEN RO/SS-A).				
GN	CALR OR CRTC.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=92013129; PubMed=1919005;				
RA	Rokeach L.A., Haselby J.A., Meilof J.F., Smeenk R.J., Unnasch T.R.,				
RA	Greene B.M., Hoch S.O.;				
RA	"Characterization of the autoantigen calreticulin.";				
RL	J. Immunol. 147:3031-3039(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=90237213; PubMed=2332496;				
RA	McCaulliffe D.P., Lux F.A., Lieu T.S., Sanz I., Hanke J., Newkirk M.M.,				
RA	Bachinski L.L., Itoh Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,				
RA	Capra J.D.;				
RT	"Molecular cloning, expression, and chromosome 19 localization of a				
RT	human Ro/SS-A autoantigen.";				
RL	J. Clin. Invest. 85:1379-1391(1990).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=92129342; PubMed=1733953;				
RA	McCaulliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;				
RT	"The 5'-flanking region of the human calreticulin gene shares				
RT	homology with the human GRP78, GRP94, and protein disulfide isomerase				
RT	promoters.";				
RL	J. Biol. Chem. 267:2557-2562(1992).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RP	Lamerdin J., McCreedy P., Stilwagen S., Ramirez M., Carrano A.;				
RP	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 18-36.				
RP	MEDLINE=92002034; PubMed=1911778;				
RA	Rojiani M.V., Finlay B.B., Gray V., Dedhar S.;				
RT	"In vitro interaction of a polypeptide homologous to human Ro/SS-A				
RT	antigen (calreticulin) with a highly conserved amino acid sequence in				
RT	the cytoplasmic domain of integrin alpha subunits.";				
RL	Biochemistry 30:9859-9866(1991).				
RN	[6]				
RP	SEQUENCE OF 18-32.				
RP	MEDLINE=90380058; PubMed=2400400;				
RA	Krause K.H., Simmerman H.K.B., Jones L.R., Campbell K.P.;				
RT	"Sequence similarity of calreticulin with a Ca2(+)-binding protein				
RT	that co-purifies with an ins(1,4,5)P3-sensitive Ca2+ store in HL-60				

cells.";
 RL Biochem. J. 270:545-548(1990).
 [7]
 RN SEQUENCE OF 18-28.
 RC
 RX TISSUE=Liver;
 RX MEDLINE=93162045; PubMed=1286669;
 RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
 RA Appel R.D., Hughes G.J.;
 RA "Human liver protein map: a reference database established by
 RT microsequencing and gel comparison.";
 RL Electrophoresis 13:992-1001(1992).
 [8]
 RN PARTIAL SEQUENCE OF 25-34; 56-62; 208-221 AND 273-278.
 RC
 RX TISSUE=Keratinocytes;
 RX MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
 RA Vandekerckhove J.;
 RA "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 [9]
 RN SEQUENCE OF 18-26.
 RC
 RX TISSUE=Colon carcinoma;
 RX MEDLINE=97295306; PubMed=9150948;
 RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
 RA "A two-dimensional gel database of human colon carcinoma proteins.";
 RL Electrophoresis 18:605-613(1997).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE THE RO AUTOANTIGEN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M84739; AA51916.1; -
 CC EMBL; M32294; AA36582.1; -
 CC EMBL; AD000092; AAB51176.1; -
 CC PIR; A37047; A37047.
 CC PIR; S11475; S11475.
 CC PIR; A42330; A42330.
 CC PIR; A46452; A46452.
 CC SWISS-2DPAGE; P27797; HUMAN.
 CC Aarhus/Ghent-2DPAGE; 9401; IEF.
 CC HSC-2DPAGE; P27797; HUMAN.
 CC MIM; 109091; -
 CC InterPro; IPR000886; -
 CC InterPro; IPR001580; -
 CC Pfam; PF00262; calreticulin; 1.
 CC PRINTS; PR00626; CALRETICULIN.
 CC PROSITE; PS00014; ER_TARGET; 1.
 CC PROSITE; PS00803; CALRETICULIN_1; 1.
 CC PROSITE; PS00804; CALRETICULIN_2; 1.
 CC PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 CC Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 KW SIGNAL 1 17
 FT CHAIN 1 17
 FT CHAIN 18 417
 FT DOMAIN 18 197
 FT DOMAIN 198 308
 FT DOMAIN 309 417
 FT DOMAIN 418 255
 FT DOMAIN 191 202
 FT REPEAT 191 202
 FT REPEAT 210 221
 FT REPEAT 227 238
 FT REPEAT 244 255
 FT REPEAT 244 255

FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
 FT REPEAT 259 269 2-1.
 FT REPEAT 273 283 2-2.
 FT REPEAT 287 297 2-3.
 FT DOMAIN 351 408 ASP/GLU/LYS-RICH.
 FT DISULFID 137 163 BY SIMILARITY.
 FT SITE 414 417 PREVENT SECRETION FROM ER.
 FT CONFLICT 35 35 MISSING (IN REF. 3).
 SQ SEQUENCE 417 AA; 48141 MW; BC37C3C0F1054FB2 CRC64;
 Query/Match 100.0%; Score 971; DB 1; Length 417;
 Best/Local Similarity 100.0%; Pred. No. 2.4e-80;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPAYFKEQFLDGDGWTSRWIESKHKSDFCGFVLSGKFGYDEKDKGLQTSQDAREYAL 60
 DB 18 EPAYFKEQFLDGDGWTSRWIESKHKSDFCGFVLSGKFGYDEKDKGLQTSQDAREYAL 77
 QY 61 SASFEPFSNKGQTLVQVQFTVKHEQNIDCGGYVKLFPSNLDQTDMDHGDSEYNIMFGPDIC 120
 DB 78 SASFEPFSNKGQTLVQVQFTVKHEQNIDCGGYVKLFPSNLDQTDMDHGDSEYNIMFGPDIC 137
 QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180
 DB 138 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 197
 RESULT 2
 CRTCL_RAT STANDARD; PRT; 416 AA.
 AC P18418; P10452;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (CALBP)
 GN CALR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=brain cortex;
 RX MEDLINE=90370496; PubMed=2395661;
 RA Murthy K.K., Banville D., Sikant C.B., Carrier F., Bell A.,
 RA Holmes C., Patel Y.C.;
 RA "Structural homology between the rat calreticulin gene product and
 RT the Onchocerca volvulus antigen Ral-1.";
 RL Nucleic Acids Res. 18:4933-4933(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=93202172; PubMed=8453984;
 RA Nakamura M., Moriya M., Baba T., Michikawa Y., Yamanobe T., Arai K.,
 RA Okinaga S., Kobayashi T.;
 RA "An endoplasmic reticulum protein, calreticulin, is transported into
 RT the acrosome of rat sperm.";
 RL Exp. Cell Res. 205:101-110(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
 RX MEDLINE=95181573; PubMed=7876339;
 RA Soennichsen B., Fuellekrug J., van Nguyen P., Diekmann W.,
 RA Robinson D.G., Mieskes G.;
 RA "Retention and retrieval: both mechanisms cooperate to maintain
 RT calreticulin in the endoplasmic reticulum.";
 RL J. Cell Sci. 107:2705-2717(1994).
 RN [4]
 RP SEQUENCE OF 270-358 FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; Latruffe N.;
 RA Lone Y.C., Bailly A., Latruffe N.;

RC TISSUE=Lung;
 RX MEDLINE=92002038; PubMed=1911780;
 RA Guan S., Falick A.M., Williams D.E., Cashman J.R.;
 RT "Evidence for complex formation between rabbit lung flavin-containing
 monooxygenase and calreticulin.";
 RL Biochemistry 30:9892-9900(1991).
 CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 LOW AFFINITY CALCIUM-BINDING SITES.
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J051138; AAA31188.1;
 DR PIR; A34154; A34154.
 DR PIR; C33208; C33208.
 DR PIR; D33208; D33208.
 DR PIR; E33208; E33208.
 DR PIR; F33208; F33208.
 DR PIR; S13046; S13046.
 DR PIR; S13047; S13047.
 DR InterPro; IPR000886;
 DR InterPro; IPR001580;
 DR Pfam; PF00626; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 418 CALRETICULIN.
 FT DOMAIN 18 197 N-DOMAIN.
 FT DOMAIN 198 308 P-DOMAIN.
 FT DOMAIN 309 418 C-DOMAIN.
 FT DOMAIN 191 235 4 X APPROXIMATE REPEATS.
 FT REPEAT 191 202 1-1.
 FT REPEAT 210 221 1-2.
 FT REPEAT 227 238 1-3.
 FT REPEAT 244 255 1-4.
 FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
 FT REPEAT 259 269 2-1.
 FT REPEAT 273 283 2-2.
 FT REPEAT 287 297 2-3.
 FT DOMAIN 351 408 ASP/GLU/LYS-RICH.
 FT DISULFID 137 163 BY SIMILARITY.
 FT SITE 415 418 PREVENT SECRETION FROM ER.
 FT VARIANT 35 35 E -> D.
 FT CONFLICT 90 90 P -> T (IN REF. 5).
 SQ SEQUENCE 418 AA; 48275 MW; B6082B689DC763A6 CRC64;

Query Match 95.9%; Score 931; DB 1; Length 418;
 Best Local Similarity 95.6%; Pred. No. 9.7e-77;
 Matches 172; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 1 EPVAYFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDQKGLQTSQDARFYAL 60
 || |||||
 Db 18 EPVAYFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDQKGLQTSQDARFYAL 77
 || |||||
 QY 61 SASFEFSSNKGQTLVVQVFTVKHEQNDICGGYVKLPNLSLDQTMHGDSEYNIMFGPDIC 120
 || |||||
 Db 78 SARFEFSSNKGQTLVVQVFTVKHEQNDICGGYVKLPNLSLDQTMHGDSEYNIMFGPDIC 137
 || |||||
 QY 121 GPCTKKVHVIFNFKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 180
 || |||||

Db 138 GPCTKKVHVIFNFKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 197
 RESULT 4
 CRTI_BOVIN
 ID CRTI_BOVIN STANDARD; PRT; 400 AA.
 AC P52193;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE CALRETICULIN, BRAIN ISOFORM 1 (CRP55) (CALREGULIN) (HACBP).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=94183174; PubMed=8135753;
 RA Matsuo K., Seta K., Yamakawa Y., Okuyama T., Shinoda T., Isobe T.;
 RT "Covalent structure of bovine brain calreticulin.";
 RL Biochem. J. 298:435-442(1994).
 CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 LOW AFFINITY CALCIUM-BINDING SITES.
 CC -!- SUBUNIT: MONOMER.
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 DR InterPro; IPR000886;
 DR InterPro; IPR001580;
 DR Pfam; PF00626; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Glycoprotein.
 FT DOMAIN 1 180 N-DOMAIN.
 FT DOMAIN 181 291 P-DOMAIN.
 FT DOMAIN 292 400 C-DOMAIN.
 FT DOMAIN 174 238 4 X APPROXIMATE REPEATS.
 FT REPEAT 174 185 1-1.
 FT REPEAT 193 204 1-2.
 FT REPEAT 210 221 1-3.
 FT REPEAT 227 238 1-4.
 FT DOMAIN 242 280 3 X APPROXIMATE REPEATS.
 FT REPEAT 242 252 2-1.
 FT REPEAT 256 266 2-2.
 FT REPEAT 270 280 2-3.
 FT DOMAIN 334 390 ASP/GLU/LYS-RICH.
 FT DISULFID 120 146 N-LINKED (GLCNAC. .).
 FT CARBOHYD 162 162 PREVENT SECRETION FROM ER (POTENTIAL).
 FT SITE 397 400
 SQ SEQUENCE 400 AA; 46381 MW; 7D4B68DFC689EEF1 CRC64;

Query Match 95.8%; Score 930; DB 1; Length 400;
 Best Local Similarity 95.0%; Pred. No. 1.1e-76;
 Matches 171; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 EPVAYFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDQKGLQTSQDARFYAL 60
 || |||||
 Db 1 DPTVYFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDQKGLQTSQDARFYAL 60
 || |||||
 QY 61 SASFEFSSNKGQTLVVQVFTVKHEQNDICGGYVKLPNLSLDQTMHGDSEYNIMFGPDIC 120
 || |||||
 Db 61 SARFEFSSNKGQTLVVQVFTVKHEQNDICGGYVKLPNLSLDQTMHGDSEYNIMFGPDIC 120
 || |||||
 QY 121 GPCTKKVHVIFNFKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 180
 || |||||
 Db 121 GPCTKKVHVIFNFKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 180
 || |||||

```

RESULT_5
CRTC_MOUSE
ID CRTC_MOUSE STANDARD; PRT: 416 AA.
AC P14211;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60).
GN CALR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.A. AND SEQUENCE OF 18-48 AND 129-161.
RX STRAIN-BALB/C; TISSUE=Liver;
RX MEDLINE=90059955; PubMed=2583110;
RA Smith M.J., Koch G.L.E.;
RT "Multiple zones in the sequence of calreticulin (CRP55, calregulin,
HACBP), a major calcium binding ER/SR protein.";
RL EMBO J. 8:3581-3586(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93013037; PubMed=1398135;
RA Mazzarella R.A., Gold P., Cunningham M., Green M.;
RT "Determination of the sequence of an expressible cDNA clone encoding
ERP60/calregulin by the use of a novel nested set method.";
RL Gene 120:217-225(1992).
RN [3]
RP SEQUENCE OF 18-38.
RX TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14926; CAA33053.1; -
DR EMBL; M92988; AAA37569.1; -
DR PIR; S06763; S06763.
DR SWISS-2DPAGE; P14211; MOUSE.
DR MGD; MGI:88252; Calr.
DR InterPro; IPR000886; -
DR InterPro; IPR001580; -
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
KW SIGNAL
FT CHAIN 1 17
FT CHAIN 18 416 CALRETICULIN.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 416 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.

```

```

FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 413 416 PREVENT SECRETION FROM ER.
SQ SEQUENCE 416 AA; 47994 MW; 24C03B00913408D8 CRC64;

Query Match 95.88; Score 930; DB 1; Length 416;
Best Local Similarity 94.4%; Pred. No. 1.2e-76;
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPAYVFEQFLDGDGWTSRWIESKHSKDFGVFLSSGKFGYDEEKDKGLQTSQDARFVAL 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 18 DPAYVFEQFLDGDGWTNRWVESKHSKDFGVFLSSGKFGYDEEKDKGLQTSQDARFVAL 77
QY 61 SASPEFSNKGQTLVQFTVKHEQNIDCGGYYVKLFPSNLSQDTMHGDSSEYNIMFGPDIC 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 78 SAKPEFSNKGQTLVQFTVKHEQNIDCGGYYVKLFPSGLDQKDMHGDSEYNIMFGPDIC 137
QY 121 GPCTKKVHVIFNYKGNVLIINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 138 GPCTKKVHVIFNYKGNVLIINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 197

RESULT_6
CRTC_DROME
ID CRTC_DROME STANDARD; PRT: 406 AA.
AC P29413; Q9VHA3;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP).
CC CRC OR CG9429.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208374; PubMed=1296819;
RA Smith M.J.;
RT "Nucleotide sequence of a Drosophila melanogaster gene encoding a
calreticulin homologue.";
RL DNA Seq. 3:247-250(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Gordon R.C., Rogers V.-H.C., Blazer R.G., Zhang Q., Chen L.X.,
RA Brandon R.C., Baxter E.G., Helt G., Nelson C.R., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Abriel J.F., Basu A., Haxendale J., Bayraktaroglu L., Baldwin D.,
RA Haller R.M., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Meson K.Y., Botchan M.R., Bouck J., Brokstein P., Brattler P.,
RA Korkova D., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Burtis K.C., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA Cherry J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA de Pablos B., Delcher A., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,

```

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Raebert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [13]
 RN SEQUENCE OF 91-124 AND 182-220.
 RX MEDLINE=90307981; PubMed=2365822;
 RA McCauliffe D.P., Zappi E., Lieu T.S., Michalak M., Sontheimer R.D.,
 RA Capra J.D.:
 RT "A human RO/SS-A autoantigen is the homologue of calreticulin and is
 RT highly homologous with onchocercal RAL-1 antigen and an alysia
 RT 'memory molecule'.";
 RL J. Clin. Invest. 86:332-335(1990).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X64461; CAA45791.1; -;
 DR EMBL; AE003683; AAF54416.1; -;
 DR PIR; A37158; A37158.
 DR FlyBase; FBgn0005585; Crc.
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER-TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 406 CALRETICULIN.
 FT REPEAT 107 107 G -> A (IN REF. 3).
 FT CONFLICT 184 184 V -> L (IN REF. 3).
 FT CONFLICT 184 184 V -> L (IN REF. 3).
 SQ SEQUENCE 406 AA; 46808 MW; 65D72C69D0BEC427 CRC64;

Query Match 72.0%; Score 699; DB 1; Length 406;
 Best Local Similarity 74.2%; Pred. No. 7.5e-56;
 Matches 132; Conservative 14; Mismatches 30; Indels 2; Gaps 2;

QY 4 VYFKEDFLDGDGWTSRWIESKHK-SDFGKFLVSSGKFGYDEEKDKLQTSQDARFYALSA 62
 DB 21 VYKNEF-DNENWEDTWIYSKPGKEFGKFLVTPGGFYNDADKDGKTSQDARFYAASR 79
 QY 63 SFEPSFNKGQTLVQVTVKHEQIDCGGYKFLFNPNSLDQTMHGDSEYNFMFGPDICGP 122
 DB 80 KFDGFSNEDKPLVQVFSVKHEQIDCGGYKFLFDCSLDQTMHGDSEYNFMFGPDICGP 139

QY 123 GTKKVVHIFNYKGNLVINKDKDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 180
 DB 140 GTKKVVHIFSYKGNHLISKDKDDVYTHFTLVIRPDNTYEVLDIDNEKVESGNLE 197
 RESULT -7
 RAL1_ONCVO STANDARD; PRT; 388 AA.
 ID RAL1_ONCVO
 AC P11012;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE RAL-1 PROTEIN PRECURSOR (41 KDA LARVAL ANTIGEN).
 GN RAL1.
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OC NCBI_Taxid=6282;
 OX [1]
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94341871; PubMed=7520419;
 RA Rokeach L.A., Zimmerman P.A., Unnasch T.R.;
 RT "Epitopes of the Onchocerca volvulus RAL1 antigen, a member of the
 RT calreticulin family of proteins, recognized by sera from patients
 RT with onchocerciasis.";
 RL Infect. Immun. 62:3696-3704(1994).
 RN [2]
 RP SEQUENCE OF 53-388 FROM N.A.
 RX MEDLINE=88273584; PubMed=2455736;
 RA Unnasch T.R., Gallin M.Y., Soboslay P.T., Ertmann K.D., Greene B.M.;
 RT "Isolation and characterization of expression cDNA clones encoding
 RT antigens of Onchocerca volvulus infective larvae.";
 RL J. Clin. Invest. 82:262-269(1988).
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M20565; AAA59056.1; -;
 DR PIR; A32507; A32507.
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Calcium-binding; Repeat; Antigen; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 388 RAL-1 PROTEIN.
 FT DOMAIN 189 253 4 X APPROXIMATE REPEATS.
 FT REPEAT 189 200 1-1.
 FT REPEAT 208 219 1-2.
 FT REPEAT 225 236 1-3.
 FT REPEAT 242 253 1-4.
 FT DOMAIN 257 295 3 X APPROXIMATE REPEATS.
 FT REPEAT 257 267 2-1.
 FT REPEAT 271 281 2-2.
 FT REPEAT 285 295 2-3.
 FT DOMAIN 353 388 ARG/LYS-RICH (BASIC).
 FT DISULFID 135 161 BY SIMILARITY.
 SQ SEQUENCE 388 AA; 45298 MW; 9537F298A2D31CD6 CRC64;

Query Match 68.4%; Score 664; DB 1; Length 388;
 Best Local Similarity 67.4%; Pred. No. 1e-52;
 Matches 120; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

QY 4 VYFKEDFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKLQTSQDARFYALSA 63

```

FT SITE      392   395                PREVENT SECRETION FROM ER.
SQ SEQUENCE  395 AA: 45616 MW;  35CA7D2EC1D56B03 CRC64;

Query Match          63.8%    Score 619.5; DB 1; Length 395;
Best Local Similarity 63.5%; Pred. No. 1.le-48;
Matches 113; Conservative 26; Mismatches 36; Indels 3; Gaps

QY      4 VYFKEQLFDGWSRSWIESKHSKDFGRKFVLSSSGKFYGDEEKGLOTSQDAERYALSAS 63
       | :|::|||:|||:|||||:|||||: |||:|:|||||:|:|:|:|:|:|:|:|:|:|:|:
DB     17 VFYEKEFNDA-SWEKRWWQSOKHKDDFGAFKLACKCFDESVDQGSIQTSDAKFFYSRAK 75

QY     64 PE-PFSNKGQTLVVOFTVKKHQNIDCGGGYVKLFPNSLDQDTMHGDGSYNFMFGPDICGP 12
       |:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB     76 FDKOFSNKGLTVIQTYYTKHEQQIDCGGGYVKVMRADADLGDFHGETPYNVFMFGPDICGP 13

QY    123 GTKKVHVIFNYKGNVLINKDIRCKDDETHLYLTILVRPNDTYEVKINDSQUESGLE 180
       |:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB    136 -TRRVHVILNYKGENKLKIETCKDSDELTHYLILNSDNTYEVKIDGESAQTGSLE 192

RESULT      9
CRQC_DICDI CRQC_DICDI STANDARD; PRT; 424 AA.
AC Q23858;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Dictyostelium discoidium (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Mueller-Taubenberger A., Gerisch G.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC !- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC !- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collabo-
CC between the Swiss Institute of Bioinformatics and the EMBL Outsta-
CC European Bioinformatics Institute. There are no restrictions on o-
CC use by non-profit institutions as long as its content is in c-
CC modified and this statement is not removed. Usage by and for comm-
CC entities requires a license agreement (See http://www.isb-sib.ch/ann-
CC or send an email to license@isb-sib.ch)
-----
EMBL: U36937; AAB87719.1; -.
DR DictyDb; DD00042; -.
DR InterPro; IPR000886; -.
DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CCHAIN 20 424 CALRETICULIN.
FT DOMAIN 191 257 4 X 12 AA APPROXIMATE REPEATS.
FT RPEAT 191 202 1-1.
FT REPEAT 211 222 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 246 257 1-4.
FT DOMAIN 260 298 3 X 11 AA APPROXIMATE REPEATS.
FT REPEAT 260 270 2-1.
FT REPEAT 274 284 2-2.
FT REPEAT 288 298 2-3.
FT DISULPID 105 137 BY SIMILARITY.
FT SITE 421 424 PREVENT SECRETION FROM ER(POTENTIAL).
```

Db 20 IYYKTF--EPDWETRWTHSTAKSDYGRFKLTSGKFYGDKAKDAGIQTSQDAKFYAISP 77
 Qy 64 F-EFPSNKGQTLVQVQFTVKHEQNTDCGGYVKLPFNSLDQTDHMGDSEYNIMFGPDICG 122
 Db 78 IASFSNEGKDLVQLQFSVKHEQDIDCGGYKLLP-SVDAAKFTGDTPYHIMFGPDICG- 135
 Qy 123 GTKVHVIFNYKGNVLNKDKDDFTHLTYTLIVRPDNTYKVIDNSQVESGSL 180
 Db 136 ATKTIHFLTYKGNLLWKKEPRCETDLSHTYTAIVIKRADRTYEVLDVQVKESGTL 193

RESULT 11
 CRIC BETVU STANDARD; PRT; 416 AA.
 ID CRIC BETVU STANDARD; PRT; 416 AA.
 AC 081919;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CALRETICULIN PRECURSOR.
 OS Beta vulgaris (Sugar beet).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
 OC Caryophyllales; Chenopodiaceae; Beta.
 OC NCBI_TaxID=3555;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=VV-D/2R5; TISSUE=Leaf;
 RA Viereck R.;
 RT "Nucleotide sequence from sugar beet calreticulin.";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)
 CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL: AJ002057; CAA05161.1; -;
 DR Mendel; 32549; Betvu; 1166; 32549.
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 DR Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 416 CALRETICULIN.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 413 416 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 416 AA; 48136 MW; 565FEC3489F77CA7 CRC64;

Query Match 54.7%; Score 531; DB 1; Length 416;
 Best Local Similarity 55.9%; Pred. No. 1.le-40;
 Matches 100; Conservative 28; Mismatches 45; Indels 6; Gaps 3;

Qy 4 VYFKEQFLDGGWTSRWIESKHSD---FGFVLSKGFYDEEKDKGLQTSQDAKFYAL 60
 Db 27 VFFDERP-EDGWPKRWVSKDWDKESMAGWBNWYTSKWNKD-ANDKGIQTSYDFRYAI 83
 Qy 61 SASPEFESNKGQTLVQVQFTVKHEQNTDCGGYVKLPFNSLDQTDHMGDSEYNIMFGPDIC 120
 Db 84 SAEPEFESNKGQTLVQVQFTVKHEQNTDCGGYVKLPFNSLDQTDHMGDSEYNIMFGPDIC 143

QY 121 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSL 179
 Db 144 GYSTKKVHAIIFYNDTHNLKIKKPCETDQLTHVYTLVIRPDATYSILIDNKEQTGSL 202

RESULT 12
 CRTC_RICCO STANDARD; PRT; 415 AA.
 ID CRTC_RICCO
 AC P93508;
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE CALRETICULIN PRECURSOR.
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97435975; PubMed=9290642;
 RA Coughlan S.J., Hastings C., Winfrey R. Jr.;
 RT "Cloning and characterization of the calreticulin gene from Ricinus
 communis L.";
 RL Plant Mol. Biol. 34:897-911(1997).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U74631; AAB71420.1; -
 CC EMBL; U74630; AAB71419.1; -
 CC Mende; 10452; Ricco;1166;10452.
 CC InterPro: IPR000886; -
 CC InterPro: IPR001580; -
 CC Pfam; PF00262; calreticulin; 1.
 CC PRINTS; PR00626; CALRETICULIN.
 CC PROSITE; PS00014; ER_TARGET; 1.
 CC PROSITE; PS00803; CALRETICULIN_1; 1.
 CC PROSITE; PS00805; CALRETICULIN_2; 1.
 CC Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 415
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 412 415 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 415 AA; 47522 MW; DD5F452E76CC7F8C CRC64;

Query Match 54.3%; Score 527; DB 1; Length 415;
 Best Local Similarity 55.3%; Pred. No. 2.4e-40;
 Matches 99; Conservative 31; Mismatches 43; Indels 6; Gaps 3;
 QY 4 VYFKEQFLDGGWTSRWIESKHKSD---FGKFVLSGKFGYGDDEKDKGLQTSQARFYAL 60
 Db 22 VFFEEF--EDGWNVWVSDWKDKDENTAGWNTYSGKNGD-PNDKGIQTSDEYRFVAI 78
 QY 61 SASPEFSNKGQTLVQFTVKHEQNDICGGYKVLFPNSLDQTDHGDSEYNMFGPDIC 120
 Db 79 SAEPEFSNKGDKTLVQFQFSVKHEQKLDCCGGYKMLSSSTDQKKFGDGTPTYSIMFGPDIC 138
 QY 121 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSL 179

Db 139 GYSTKKVHAILNYNDTHNLKIKKPCETDQLTHVYTLVIRPDATYSILIDNKEQTGSL 197

RESULT 13
 CRTC_PUAR STANDARD; PRT; 421 AA.
 ID CRTC_PUAR
 AC Q9XF98;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CALRETICULIN PRECURSOR.
 OS Prunus armeniaca (Apricot).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Rosales; Rosaceae; Prunus.
 OX NCBI_TaxID=36596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BERGERON; TISSUE=Mesocarp, and Endocarp;
 RA Mdeguite-A-Mdeguite D., Fils-Lycaon B.R.;
 RT "Molecular cloning and nucleotide sequence of a calreticulin from
 RT apricot (Prunus armeniaca cv. Bergeron).";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF134733; AAD32207.1; -
 CC InterPro: IPR000886; -
 CC InterPro: IPR001580; -
 CC Pfam; PF00262; calreticulin; 1.
 CC PRINTS; PR00626; CALRETICULIN.
 CC PROSITE; PS00014; ER_TARGET; 1.
 CC PROSITE; PS00803; CALRETICULIN_1; 1.
 CC PROSITE; PS00804; CALRETICULIN_2; 1.
 CC PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 CC Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
 FT SIGNAL 1 22
 FT CHAIN 23 421
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 418 421 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 421 AA; 48416 MW; 4F5F94CBAAG6C690 CRC64;

Query Match 54.3%; Score 527; DB 1; Length 421;
 Best Local Similarity 55.3%; Pred. No. 2.5e-40;
 Matches 99; Conservative 28; Mismatches 46; Indels 6; Gaps 3;
 QY 4 VYFKEQFLDGGWTSRWIESKHKSD---FGKFVLSGKFGYGDDEKDKGLQTSQARFYAL 60
 Db 26 VFFEEF--EDGDKRWVTSWKKNLAGEWNTYSGKNGD-PNDKGIQTSDEYRFVAI 82
 QY 51 SASPEFSNKGQTLVQFTVKHEQNDICGGYKVLFPNSLDQTDHGDSEYNMFGPDIC 120
 Db 83 SAEPEFSNKGDKTLVQFQFSVKHEQKLDCCGGYKMLSSSTDQKKFGDGTPTYSIMFGPDIC 142
 QY 121 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSL 179
 Db 143 GYSTKKVHAILNYNDTHNLKIKKPCETDQLTHVYTLVIRPDATYSILIDNKEQTGSL 201
 RESULT 14
 CRTC_ORISA

ID	CRTC_ORVSA	STANDARD;	PRT;	424 AA.
Q9SLI8;				
AC	01-OCT-2000	(Rel. 40, Created)		
DT	01-OCT-2000	(Rel. 40, Last sequence update)		
DT	01-OCT-2000	(Rel. 40, Last annotation update)		
DT	CALRETICULIN PRECURSOR.			
OS	Oriza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;			
OC	Oryza.			
OC	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Komatsu S., Li Z.;			
RT	"Cloning the cDNA encoding a calcium-binding protein which involved in			
RT	the regeneration of rice cultured suspension cells.";			
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.			
CC	-I- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND			
CC	LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).			
CC	-I- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).			
CC	-I- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AB021259; BAA88900.1; -			
DR	InterPro; IPR000886; -			
DR	InterPro; IPR001580; -			
DR	Pfam; PF00262; calreticulin. 1.			
DR	PRINTS; PR00626; CALRETICULIN.			
DR	PROSITE; PS00014; ER_TARGET; 1.			
DR	PROSITE; PS00803; CALRETICULIN_1; 1.			
DR	PROSITE; PS00804; CALRETICULIN_2; 1.			
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 1.			
KW	Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.			
FT	SIGNAL	1	29	POTENTIAL.
FT	CHAIN	30	424	CALRETICULIN.
FT	CARBOHYD	61	61	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SITE	421	424	PREVENT SECRETION FROM ER (POTENTIAL).
SQ	SEQUENCE	424 AA;	47925 MW;	3173634FI1580FC14 CRC64;

Query Match	54.3%	Score 527:	DB 1:	Length 424:
Best Local Similarity	55.9%	Pred. No. 2.5e+40:		
Matches 100:	Conservative 31:	Mismatches 42:	Indels 6:	Gaps 3:
QY	4	VYKEQFLLDGGGTSRWIESKHKSD----	FCKFVLLSKCFYCDEEKDKGLQTSQDARYAL	60
DB	31	VFFQKEF--EDGESRWVKSEWKDENMAGENNHTSGKNGDPE-DKGIOTSEYRIFYAI	87	
QY	61	SASFEPFSNGQTLLVQFTVYKHEQIDCGGYVKLFPSNLDQTDHMGDSEYNIIMFGPDIC	120	
DB	88	SAEYFESNKKDKTLVLQFSVKYHEQLDGGYVKLLGGVDVQKFKGGDTPYPSIMFGPDIC	147	
QY	121	PGPTKKVHVITFNKYGKNVLINKDIRCKDDBEFTHTYTLIVRPDNTYEVKINDNSQVESGSL	179	
DB	148	GYSTKKVHTITFTNRDNKHLIAKKDVPCEDTQLSHVYTLIIHPDAPYSILINDNVEKQSGSI	206	

RESULT 15		
ID	CRT2_BOVIN	STANDARD; PRT; 421 AA.
AC	P42918;	
DT	01-NOV-1995	(Rel. 32, Created)
DT	01-NOV-1995	(Rel. 32, Last sequence update)
DT	15-JUL-1998	(Rel. 36, Last annotation update)
DE	CALRETICULIN, BRAIN ISOFORM 2	PRECURSOR (CRP55)
OS	Bos taurus (Bovine).	(CALREGULIN) (HACBP).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RN SEQUENCE FROM N.A.
RP
RC TISSUE=Brain;
RC MEDLINE=93385184; Pubmed=8373827;
RA Liu N., Fine R.E., Johnson R.J.;
RA "Comparison of cDNAs from bovine brain coding for two isoforms of
RT calreticulin.";
RL Biochim. Biophys. Acta 1202:70-76(1993).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC -1- LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; LJ3462; AAC37307.1; -
DR InterPro; IPR000886; -
DR InterPro; IPR001580; -
DR Pfam; PF00262; calreticulin.1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 34
FT CHAIN 35 421
FT DOMAIN 35 201
FT DOMAIN 35 201
FT DOMAIN 302 312
FT DOMAIN 313 421
FT DOMAIN 195 259
FT REPEAT 195 206
FT REPEAT 214 225
FT REPEAT 231 242
FT REPEAT 248 259
FT DOMAIN 263 301
FT REPEAT 263 273
FT REPEAT 277 287
FT REPEAT 291 301
FT DOMAIN 366 411
FT DISULFID 141 167
FT CARBOHYD 183 183
FT SITE 418 421
SQ SEQUENCE 421 AA; 48812 MW; 0257E959F71528BC CRC64;
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
PREVENT SECRETION FROM ER.
0257E959F71528BC CRC64;

```

Query Match      53.8%; Score 522.5; DB 1; Length 421;
Best Local Similarity 55.6%; Pred. No. 6.3e-40;
Matches 109; Conservative 12; Mismatches 30; Indels 45; Gaps 4;

QY      1  EPAYVFKEQFLDGDGWTSRWIESKHKSD-----FGFVLSCKKFFGDEE 44
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       35  EETVFFSEQFLT-----LDLUYKRAKSLSTREALSMKVGIENFCFSEISFLQESI 86
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY      45  KDKGLQPTSQDAREFVALSASPEPSNKGQTLLVQFTVKEQNICDGGGYVKVFLNSLDQTD 104
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       87  KSHGRRT-----LVGCSPWGHE-----EQNICDGGGYVNVFFAGLDQTD 125
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY     105  MHGDSYNIIMFGPDICGPGTKKVHVIFNYKGKNVLINKDIRCKDDEFFHLYTLVLRPDNT 164
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     126  MHGDSYNIIMFGPDICGPGTKKVHVIFNYKGKNVLINKDIRCKDDEFFHLYTLVLRPNNT 185
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:10:07 ; Search time 92.24 seconds
(without alignments)
258.184 Million cell updates/sec

Title: US-09-807-148-4
Perfect score: 971
Sequence: 1 EPAYVFKKQFLDGDGWTSRW.....PDNTYEKIDNSQVESGSLE 180

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_unclassified:*
14: sp_vertebrate:*
15: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835	86.0	411	13	Q91710
2	804	82.8	405	5	Q26268
3	797	82.1	343	13	Q91711
4	796	82.0	419	13	Q98984
5	785	80.8	417	13	Q9PUC1
6	759	78.2	410	5	Q16893
7	731	75.3	421	5	Q9U850
8	699	72.0	406	5	Q9U916
9	681.5	70.2	403	5	Q76961
10	680	69.0	387	5	Q97372
11	673	69.3	375	5	Q18478
12	539	55.5	412	10	Q40040
13	539	55.5	415	10	Q40041
14	536	55.2	421	10	Q43712
15	517	53.2	389	10	Q40567
16	514.5	53.0	427	10	Q9FYV2
17	507.5	52.3	396	5	Q45034
18	477.5	49.2	350	5	Q26514
19	464.5	47.8	240	10	Q9ST29

20	460	47.4	318	13	Q9PTX7
21	450	46.3	214	4	Q9UDG2
22	448.5	46.2	403	5	Q9XYF8
23	446.5	46.0	401	5	Q9U9N9
24	409.5	42.2	291	5	Q9Y1V1
25	392	40.4	397	5	Q94592
26	373	38.4	321	13	Q9U5G0
27	347	35.7	422	10	Q22502
28	329.5	33.9	321	10	Q41799
29	302.5	31.2	101	10	Q40751
30	289.5	29.8	582	3	Q9HFC6
31	264.5	27.2	559	5	Q9NG26
32	262	27.0	622	13	Q98985
33	251	25.8	70	10	Q9SXW3
34	248	25.5	582	5	Q04702
35	240	24.7	583	5	Q9VXF6
36	238.5	24.6	582	5	Q76214
37	238.5	24.6	582	5	Q9TVF3
38	231.5	23.8	272	4	Q16094
39	224	23.1	543	5	Q9VYP4
40	224	23.1	556	5	Q9I7S9
41	221.5	22.8	605	5	Q02393
42	207	21.3	545	5	Q9VAL7
43	200.5	20.6	532	10	Q9LY26
44	131	13.5	68	6	Q9TS71
45	129	13.3	297	5	Q24097

ALIGNMENTS

RESULT 1

ID Q91710 PRELIMINARY; PRT; 411 AA.
AC Q91710;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CALRETICULIN PRECURSOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Anphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=93074997; PubMed=1445218;
RA Treves S., Forzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous system";
RL Biochem. J. 287:579-581(1992).
DR EMBL: X67597; CAA47866.1; -;
DR InterPro: IPR000886; -;
DR InterPro: IPR001580; -;
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 12 POTENTIAL.
FT CHAIN 13 411 CALRETICULIN.
SQ SEQUENCE 411 AA; 48344 MW; 891DA66E00EBEFA CRC64;

Query Match 86.0%; Score 835; DB 13; Length 411;

Best Local Similarity 83.9%; Pred. No. 5.7e-71;

Matches 151; Conservative 15; Mismatches 14; Indels 0; Gaps 0;

QY 1 EPAYVFKKQFLDGDGWTSRWTSKHKHSDFGKFLSSGKFGDEEKGLOTQSQDARFVAL 60

OX	NCBI_TaxID=8355;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BRAIN;
RX	MEDLINE=93074997; PubMed=1445218;
RA	Treves S., Zorzato F., Pozzan T.;
RT	*Identification of calreticulin isoforms in the central nervous system.*;
RL	Biochem. J. 287:579-581(1992).
DR	EMBL; X67598; CAA47867.1; -.
DR	InterPro; IPR001580; -.
DR	Pfam; PF00262; calreticulin; 1.
DR	PRINTS; PR00626; CALRETICULIN.
DR	PROSITE; PS00803; CALRETICULIN_1; 1.
DR	PROSITE; PS00804; CALRETICULIN_2; 1.
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
FT	NON_TER 1
FT	NON_TER 343 343
SQ	SEQUENCE 343 AA; 40105 MW; 3E7DDAFA33B91DE1 CRC64;
Query Match 82.1%; Score 797; DB 13; Length 343;	
Best Local Similarity 85.2%; Pred. No. 1.8e-67;	
Matches 144; Conservative 12; Mismatches 13; Indels 0; Gaps	
QY	12 DGDGWTSRWTESKHKDFGKFVLSGGKYGDEKDKGLQTSQDARYALASFEPSNKG 71
Db	1 DGDGTQRWWESKHSYDGKFKUSAGFYGDSEKDKGLQTSQDARYAMSSRFESFNKD 60
QY	72 QTLVVOFTVKHEQNIDCGGYVKLFNSLDQTDMHGDSYNNMFGPDICGPGTKKVHVIF 131
Db	61 QTLVIQFSVKHEQNIDCGGYVKLFPADLEQTEHGESEYNMFGPDICGPGTKKVHVIF 120
QY	132 NYGKNVLNKDIRCKDDETHLYTLIVRPDNTYEVKIDNQSVESGLE 180 :
Db	121 QYKKKLQINKDIRCKDDSFTHLYTLIVRPDNTYEVKIDNSKVESGLE 169 :
RESULT	4
Q98984	
ID	Q98984 PRELIMINARY; PRT; 419 AA.
AC	Q98984;
DT	01-FEB-1997 (TREMBLrel. 02, Created)
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE	CALRETICULIN.
OS	Rana rugosa (Frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX	NCBI_TaxID=8410;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=96234004; PubMed=8654561;
RA	Yamamoto S., Nakamura M.;
RT	*calnexin: its molecular cloning and expression in the liver of the frog, Rana rugosa.*;
RL	FEBS Lett. 387:27-32(1996).
DR	[2]
RP	SEQUENCE FROM N.A.
RA	Yamamoto S.;
RT	*strong expression of the calreticulin gene in the liver of Rana rugosa tadpoles.*;
RL	J. Exp. Zool. 0:0-0(1996).
DR	EMBL; D78589; BAA11425.1; -.
DR	InterPro; IPR000886; -.
DR	InterPro; IPR001580; -.
DR	Pfam; PF00262; calreticulin; 1.
DR	PRINTS; PR00626; CALRETICULIN.
DR	Prodom; PD001866; -; 1.
DR	PROSITE; PS00803; CALRETICULIN_1; 1.
DR	PROSITE; PS00804; CALRETICULIN_2; 1.
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR	PROSITE; PS00014; FR.TARCF.T.UNKNOWN.1

DT	01-JAN-1999 (TREMBlrel. 09, Last sequence update)
DE	01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE	CALRETICULIN.
GN	CRT-1.
OS	Amblyomma americanum.
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC	-Parasitiformes; Ixodida; Ixodidae; Amblyomma.
NCBI_TaxID=6943;	
[1]	
RN	SEQUENCE OF 49-410 FROM N.A.
RP	TISSUE-SALIVARY GLANDS;
RC	Jaworski D.C.; Simmen F.A., Lamoreaux W.J., Coons L.B., Muller M.T.,
RA	Needham G.R.;
RA	"A secreted calreticulin protein in Ixodid tick (<i>Amblyomma americanum</i>)
RT	saliva.";
RT	J. Insect Physiol. 41:369-375(1995).
RL	[2]
RN	SEQUENCE FROM N.A.
RP	TISSUE-SALIVARY GLANDS;
RC	Jaworski D.C.;
RA	Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RL	[3]
RN	SEQUENCE FROM N.A.
RP	TISSUE-SALIVARY GLANDS;
RC	Fain-Thornton J.M., Jaworski D.C., Needham G.R.;
RA	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RL	EMBL: U07708; AAC79094.1; -.
DR	InterPro; IPRO00886; -.
DR	InterPro; IPRO01380; -.
DR	Pfam; PF00262; calreticulin; 1.
DR	.PRINTS; P00626; CALRETICULIN.
DR	ProDom; PD001866; -; 1.
DR	PROSITE; PS00803; CALRETICULIN_1; 1.
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR	PROSITE; PS00014; ER-TARGET; UNKNOWN_1.
SQ	SEQUENCE 410 AA; 47485 MW; 32CCB8750A17DC34 CRC64;
Query Match 78.2%; Score 759; DB 5; Length 410;	
Best Local Similarity 76.7%; Pred.No. 8.6e-64;	
Matches 138; Conservative 15; Mismatches 27; Indels 0; Gaps	
QY	1 EPAVFYKEQFLDGDGTSRWIESKHKSDFGKFVLSSGFYGDDEKDKGLQTSQDAIFYAL 60 : : :- : : :- : :- :
Db	17 DPTVFYKEEFDGDMGKSRWVESTKGDNLGKFLVSAGRYGDAEKSKGLQTSEDARFYGI 76
QY	61 SASPEFPFSNGOTLVVVQFTVKHEQNIDCGGVYKFLPNSLDQTMHGSSEYNIMPGPDIC 120 : : :- :- :- :- :- :
Db	77 SSKFEFNEGTLLVVQFTVKHEQNIDCGGVYKFLDQLDSRLHGSPYKIMPPDIC 136
QY	121 GPCTKKVHVIFNYKKNVLINKDIRCKDEFTHLYTLVRPNTVEYKIDNSOVESGLE 180 : : :- :- :- :- :- :
Db	137 GPCTKKVHVIFNYKKNHLINKIEICKDDVFTHLYTLVLRPNTVYKIDNEVAEKGELE 196
RESULT 7	
IDQUSC	PRELIMINARY; PRT; 421 AA.
AC	Q9U6S0
AD	Q9U6S0;
DT	01-MAY-2000 (TREMBlrel. 13, Created)
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE	01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE	CALRETICULIN PRECURSOR.
GN	CALRET.
OS	Strongylocentrotus purpuratus (Purple sea urchin).
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC	Echinoidea; Euchinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC	Strongylocentrotus.
OX	NCBI_TaxID=7668;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	Susan J.M., Just M.L., Lennarz W.J.;
RT	"Cloning and Characterization of Alphas Integrin and Calreticulin in

OS	Dirofi
OC	Eukary
OC	Onchoc
OY	NCPT


```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-99094497; PubMed-9879888;
RT  "Molecular characterization of a calcium-binding protein from the
RL  filarial parasite Dirofilaria immitis."
RT  Mol. Biochem. Parasitol. 97:69-79(1998).
DR  EMBL: AF052978; AAD03405.1;
DR  InterPro: IPR001580;
DR  Pfam: PF00262; calreticulin; 1.
DR  PRINTS: PR00626; CALRETICULIN.
DR  ProDom: PD001866; 1.
DR  PROSITE: PS00803; CALRETICULIN_1; 1.
DR  PROSITE: PS00804; CALRETICULIN_2; 1.
DR  PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW  Signal.
FT  SIGNAL.
FT  CHAIN
FT  CHAIN
SQ  SEQUENCE 387 AA; 44941 MW; E7741BF6AAFA5885 CRC64;

Query Match 70.0%; Score 680; DB 5; Length 387;
Best Local Similarity 68.5%; Pred. No. 2.3e-56;
Matches 122; Conservative 27; Mismatches 27; Indels 2; Gaps 2;

QY  4 VYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYALSAS 63
DB  19 IYFKEEFSDDWENRWIKSKHKDDFGKWEISHGKFGYGDVAKDKGLTKTQDAKFYSIGAK 77
QY  64 FEP-FSNKGOTLVQFTVKHEQNDICGGYVVKLPFNSLDQTDHMGDSEYNIMFGPDICGP 122
DB  78 FDKGFSNKGKSLVQFSVKHEQDIDCGGGYVVKLMSNNVLEDFHGETPYHIMFGPDICGP 137
QY  123 GTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 180
DB  138 GTKKVHVIFHYKGNHMIKKDIRCKDDVFTHLTYLIVNSDNTYEVQIDGKAESGELE 195
QY  123 GTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 180
DB  138 GTKKVHVIFHYKGNHMIKKDIRCKDDVFTHLTYLIVNSDNTYEVQIDGKAESGELE 195

RESULT 11
ID  O18478 PRELIMINARY; PRT; 375 AA.
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT  01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE  RAL-1 PROTEIN (FRAGMENT).
OS  Litomosoides sigmodontis.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC  Onchocercidae; Litomosoides.
OX  NCBI_TaxID=42156;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MacLennan K., Hoffman W.H., Taylor D.W.;
RL  Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AJ001621; CAA04877.1;
DR  InterPro: IPR001580;
DR  Pfam: PF00262; calreticulin; 1.
DR  PRINTS: PR00626; CALRETICULIN.
DR  PROSITE: PS00803; CALRETICULIN_1; 1.
DR  PROSITE: PS00804; CALRETICULIN_2; 1.
DR  PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW  Signal.
FT  SIGNAL.
FT  CHAIN
FT  CHAIN
SQ  SEQUENCE 375 AA; 43642 MW; 03F7642FBFF7A5B8 CRC64;

Query Match 69.3%; Score 673; DB 5; Length 375;
Best Local Similarity 68.5%; Pred. No. 1e-55;
Matches 122; Conservative 26; Mismatches 28; Indels 2; Gaps 2;

QY  4 VYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYALSAS 63
DB  19 VYFKEEFLDDEWENRWIKSKHKDDFGKWEISHGKFGYGDVAKDKGLTKTQDAKFYSIGAK 77

```

```

QY  64 FEP-FSNKGOTLVQFTVKHEQNDICGGYVVKLPFNSLDQTDHMGDSEYNIMFGPDICGP 122
DB  78 FDKGFSNKGKSLVQFSVKHEQDIDCGGGYVVKLMSANLEDFHGETPYHIMFGPDICGP 137
QY  123 GTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 180
DB  138 GTKKVHVIFHYKGNHMIKKDIRCKDDVFTHLTYLIVNSDNTYEVQIDGKAESGELE 195

RESULT 12
ID  Q40040 PRELIMINARY; PRT; 412 AA.
AC  Q40040;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE  CALRETICULIN (FRAGMENT).
GN  CRH1.
OS  Hordeum vulgare (Barley).
OC  Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC  Magnoliophyta; Liliopsida; Poales; Poaceae; Triticeae;
OC  Hordeum.
OX  NCBI_TaxID=4513;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-MOREX; TISSUE-OVARY;
RX  MEDLINE-94339696; PubMed-7914763;
RA  Chen F., Hayes P.M., Mulroony D., Pan A.;
RT  "Identification and characterization of cDNA clones encoding plant
RT  calreticulin in barley."
RL  Plant Cell 6:835-843(1994).
DR  EMBL: L27348; AAA32948.1;
DR  Mendel; 8546; Horvu; 1166; 8546.
DR  InterPro: IPR000886;
DR  InterPro: IPR001580;
DR  Pfam: PF00262; calreticulin; 1.
DR  PRINTS: PR00626; CALRETICULIN.
DR  PROSITE: PS00803; CALRETICULIN_1; 1.
DR  PROSITE: PS00804; CALRETICULIN_2; 1.
DR  PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
DR  PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW  Calcium-binding.
FT  NON_TER 1
FT  NON_TER 412
SQ  SEQUENCE 412 AA; 47037 MW; 37F6C95D6AA78AB0 CRC64;

Query Match 55.5%; Score 539; DB 10; Length 412;
Best Local Similarity 56.4%; Pred. No. 5.3e-43;
Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY  4 VYFKEQFLDGDGWTSRWIESKHKSD---FGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60
DB  20 VYFQEKF--EDGWESRWVSKSEKKDENMAGNWNHTSGKWHGDAE-DKGIQTSYDFRYAI 76
QY  61 SASPEFSNKGOTLVQFTVKHEQNDICGGYVVKLPFNSLDQTDHMGDSEYNIMFGPDIC 120
DB  77 SAEYFESNKKDITLVQFTVKHEQNDICGGYVVKLGGVDQKKRPGGDTGYIMFGPDIC 136
QY  121 GPGTKKHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 179
DB  137 GYSTKKVHTILTCKNGKNHLIKKDVPCETDQLSHVYTLIIRPDATYSILIDNEEKOTGSI 195

RESULT 13
ID  Q40041 PRELIMINARY; PRT; 415 AA.
AC  Q40041;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE  CALRETICULIN (FRAGMENT).
GN  CRH2.

```

OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticeae;
 OC Hordeum
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MOREX; TISSUE-OVARY;
 RX MEDLINE=94339696; PubMed=7914763;
 RA Chen F., Hayes P.M., Mulrooney D., Pan A.;
 RT "Identification and characterization of cDNA clones encoding plant
 calreticulin in barley."
 RL Plant Cell 6:835-843(1994).
 DR EMBL; L27349; AAA32949.1; -
 DR Mendel; 8547; Horvu; 1166; 8547.
 DR InterPro; IPR000886; -
 DR InterPro; IPR001580; -
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Calcium-binding.
 FT SIGNAL 1 1
 FT NON_TER 415 415
 FT SEQUENCE 415 AA; 47359 MW; 2897914812FBE33E CRC64;

Query Match 55.5%; Score 539; DB 10; Length 415;
 Best Local Similarity 56.4%; Pred. No. 5.3e-43;
 Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHSD---FGKFLVSSGKFGYGDDEKDKGLQTSODARFYAL 60
 Db 23 VFFQERF--EDGWESRWKSEWKNMAGENNHTSGKWNQDAE-DKGIQTSEDYRFYAI 79
 QY 61 SASFEPSNKGOTLVVQFTVKHEQNDICGCGYVVKLPNSLDQDMHGDSEYNIMFGPDIC 120
 Db 80 SAEPFESNKDKTLVLQFTVKHEQKLDCCGGYVVKLGDDVDQKKFGDTPYIGIMFGPDIC 139
 QY 121 GPCTKKVHVFYFNKGNVLINKDIRCKDDEFTLTYLIVRPDNTYEVKIDNSOVESGSL 179
 Db 140 GYSTKKVHTLTNGKNHLLIKKDVPCETQLDTHVYTLIRPDATYSILIDNEEKQTGSI 198

RESULT 14
 Q43712 ID Q43712 PRELIMINARY; PRT; 421 AA.
 AC Q43712;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CALCIUM-BINDING PROTEIN PRECURSOR.
 GN CRT1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
 OC Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VAR MERIT; TISSUE-ROOT TIP;
 RA Napier R.M., Trueman S., Henderson J., Boyce J.M., Hawes C.R.,
 RA Fricker M.D., Venis M.A.;
 RL J. Exp. Bot. 46:1603-1613(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96309381; PubMed=8704156;
 RA Dresselhaus T., Hagel C., Loerz H., Kranz E.;
 RT "Isolation of a full-length cDNA encoding calreticulin from a PCR
 library of in vitro zygotes of maize."
 RL Plant Mol. Biol. 31:23-34(1996).

DR EMBL; Z46772; CAA86728.1; -
 DR EMBL; X89813; CAA61939.1; -
 DR Mendel; 11228; Zeama; 1166; 11228.
 DR InterPro; IPR000886; -
 DR InterPro; IPR001580; -
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PRODOM; PD001866; -; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Signal; Calcium-binding.
 FT SIGNAL 1 25
 FT CHAIN 26 421
 FT CHAIN 26 421
 FT SEQUENCE 421 AA; 48012 MW; 5AAE02B77ED3126D CRC64;

Query Match 55.2%; Score 536; DB 10; Length 421;
 Best Local Similarity 56.4%; Pred. No. 1e-42;
 Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHSD---FGKFLVSSGKFGYGDDEKDKGLQTSODARFYAL 60
 Db 27 VFFQERF--EDGWESRWKSEWKNMAGENNHTSGKWNQDAE-DKGIQTSEDYRFYAI 83
 QY 61 SASFEPSNKGOTLVVQFTVKHEQNDICGCGYVVKLPNSLDQDMHGDSEYNIMFGPDIC 120
 Db 84 SAEPFESNKDKTLVLQFTVKHEQKLDCCGGYVVKLGDDVDQKKFGDTPYIGIMFGPDIC 143
 QY 121 GPCTKKVHVFYFNKGNVLINKDIRCKDDEFTLTYLIVRPDNTYEVKIDNSOVESGSL 179
 Db 144 GYSTKKVHTLTNGKNHLLIKKDVPCETQLDTHVYTLIRPDATYSILIDNEEKQTGSI 202

RESULT 15
 Q40567 ID Q40567 PRELIMINARY; PRT; 389 AA.
 AC Q40567;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CALRETICULIN (FRAGMENT).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95290993; PubMed=7773014;
 RA Dencke J., Carlsson L.E., Vidal S., Hoglund A.S., Ek B.,
 RA van Zeijl M.J., Sinjorgo K.M., Palva E.T.;
 RT "The tobacco homolog of mammalian calreticulin is present in protein
 complexes in vivo."
 RL Plant Cell 7:391-406(1995).
 DR EMBL; X85382; CAA59694.1; -
 DR Mendel; 9293; Nicta; 1166; 9293.
 DR InterPro; IPR000886; -
 DR InterPro; IPR001580; -
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00803; CALRETICULIN_1; UNKNOWN_1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 FT NON_TER 1 1
 FT SEQUENCE 389 AA; 44565 MW; F765C036E395EFDE CRC64;

Query Match 53.2%; Score 517; DB 10; Length 389;
 Best Local Similarity 54.7%; Pred. No. 5.9e-41;
 Matches 98; Conservative 28; Mismatches 47; Indels 6; Gaps 3;

Search completed: October 21, 2001, 03:10:08
Job time: 343 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:05:46 ; Search time 76 Seconds
(without alignments)
48.659 Million cell updates/sec

Title: US-09-807-148-5
Perfect score: 327
Sequence: 1 CGPCFKVHVIFNYKKNVL.....PDNTVEVKIDNSQVSGSLE 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	61	21	AA192352
2	327	100.0	180	21	AA192351
3	327	100.0	400	21	AA192350
4	327	100.0	401	18	AA111156
5	327	100.0	417	10	AA192276
6	327	100.0	417	20	AA100927
7	327	100.0	417	21	AA192349
8	318	97.2	60	21	AA192354
9	318	97.2	280	21	AA192355
10	280	85.6	403	17	AA104171
11	258	78.9	49	21	AA192353

12	251	76.8	336	12	AA123112	Partial sequence o
13	190	58.1	122	20	AA100924	Human cClqR binding
14	190	58.1	122	20	AA100926	Rat cClqR binding
15	186	56.9	415	22	AA166341	Castor bean calret
16	186	56.9	415	22	AA166343	Castor bean calret
17	184	56.3	385	21	AA132385	Human secreted pro
18	182	55.7	122	20	AA100925	Mouse cClqR bindin
19	178	54.4	312	21	AA124609	Arabidopsis thalia
20	178	54.4	312	21	AA147933	Arabidopsis thalia
21	178	54.4	332	21	AA130998	Arabidopsis thalia
22	178	54.4	421	21	AA124608	Arabidopsis thalia
23	178	54.4	421	21	AA147932	Arabidopsis thalia
24	178	54.4	424	21	AA124607	Arabidopsis thalia
25	178	54.4	424	21	AA147931	Arabidopsis thalia
26	178	54.4	441	21	AA130997	Arabidopsis thalia
27	178	54.4	444	21	AA130996	Arabidopsis thalia
28	115.5	35.3	593	16	AA171094	Calnexin sequence.
29	100	30.6	417	21	AA177953	A. thaliana enviro
30	85.5	26.1	394	21	AA126285	Arabidopsis thalia
31	85.5	26.1	394	21	AA146612	Arabidopsis thalia
32	85.5	26.1	530	21	AA126284	Arabidopsis thalia
33	85.5	26.1	530	21	AA146611	Arabidopsis thalia
34	85.5	26.1	567	21	AA146610	Arabidopsis thalia
35	82.5	25.2	542	22	AA166342	Castor bean calnex
36	77.5	23.7	394	21	AA104449	Arabidopsis thalia
37	77.5	23.7	532	21	AA104448	Arabidopsis thalia
38	77.5	23.7	548	21	AA104447	Arabidopsis thalia
39	61.5	18.8	337	22	AA145695	P. falciparum yfgb
40	61	18.7	868	20	AA137731	Protein involved i
41	59.5	18.2	220	21	AA153217	Bridge-1 homologou
42	59	18.0	426	21	AA157656	Arabidopsis thalia
43	59	18.0	1844	21	AA181250	Plasmodium faicipa
44	58.5	17.9	1188	21	AA18183	Plasmodium faicipa
45	56.5	17.3	332	21	AA102857	Human G protein co

ALIGNMENTS

RESULT 1

AA192352

ID AA192352 standard; Protein; 61 AA.

XX AA192352;

AC AA192352;

DT 10-AUG-2000 (first entry)

XX 10-AUG-2000 (first entry)

XX Recombinant human calreticulin residues 120-180.

DE MBP-calreticulin; maltose binding protein; angioogenesis; inhibition;

XX endotheial cell; anti-angiogenic; neuroprotective; antidiabetic;

KW cytosolic; dermalogical; immunosuppressive; anti-inflammatory; hepatic;

KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.

OS Synthetic.

XX WO2000020577-A1.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.

XX 05-OCT-1998; 98US-0103438.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Tosato G, Pike SE, Yao L;

XX WPI; 2000-303767/26.

XX Inhibiting endothelial cell growth and angioogenesis using calreticulin,

XX useful for suppressing tumor growth

PT

PS Claim 4; Page 82-83; 99pp; English.

XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).

XX Sequence 61 AA;
 SQ

Query Match 100.0%; Score 327; DB 21; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.4e-38;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60
 |||||
 Db 1 CGPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60
 |||||

QY 61 E 61
 |
 Db 61 e 61

RESULT 2
 AAY92351
 ID AAY92351 standard; Protein; 180 AA.
 AC AAY92351;
 DT 10-AUG-2000 (first entry)
 XX Human vasostatin (calreticulin N-terminal 180 amino acids).
 DE MBP-calreticulin; maltose binding protein; vasostatin; N-terminal;
 KW angiogenesis; inhibition; endothelial cell; anti-angiogenic;
 KW neuroprotective; antidiabetic; cytostatic; dermatological; hepatic;
 KW immunosuppressive; antiinflammatory; anti-atherosclerotic;
 KW gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.
 OS Synthetic.
 XX WO200020577-A1.
 XX 13-APR-2000.
 XX 05-OCT-1999; 99WO-US23240.
 XX 06-OCT-1998; 98US-0103438.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Tosato G, Pike SE, Yao L;
 XX WPI; 2000-303767/26.
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth

PS Claim 4; Page 82; 99pp; English.

XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).

XX Sequence 180 AA;
 SQ

Query Match 100.0%; Score 327; DB 21; Length 180;
 Best Local Similarity 100.0%; Pred. No. 5.6e-38;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60
 |||||
 Db 120 CGPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 179
 |||||

QY 61 E 61
 |
 Db 180 e 180

RESULT 3
 AAY92350
 ID AAY92350 standard; Protein; 400 AA.
 AC AAY92350;
 XX 10-AUG-2000 (first entry)
 XX Recombinant human MBP-calreticulin.
 DE MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.
 OS WO200020577-A1.
 XX 13-APR-2000.
 XX 05-OCT-1999; 99WO-US23240.
 XX 06-OCT-1998; 98US-0103438.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Tosato G, Pike SE, Yao L;
 XX WPI; 2000-303767/26.
 XX N-PSDB; AAA09346, AAA09347.
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth

PS Claim 4; Page 80-81; 99pp; English.
 XX Recombinant human MBP-calreticulin comprises the sequence of human
 CC calreticulin (see AAY92349) without the 17 N-terminal amino acids.

CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The
 CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).

XX SQ Sequence 400 AA;

Query Match 100.0%; Score 327; DB 21; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1.6e-37;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
 Db 120 cpgptkkvkvifnykgnvlinkdirckddefthlytlivrpdpntyevkidnsqvesgsl 179
 QY 61 E 61
 Db 180 e 180

RESULT 4

ID AAW11156
 AC AAW11156 standard; peptide; 401 AA.

XX 31-MAY-1997 (first entry)

XX Calreticulin.

XX calreticulin; C-domain; restenosis; inhibitor.

XX Homo sapiens.

XX WO9636643-A1.

XX 21-NOV-1996.

XX 17-MAY-1996; 96WO-IB00471.

XX 16-MAY-1996; 96US-0649417.

XX 17-MAY-1995; 95US-0442844.

XX (UYAL-) UNIV ALBERTA.

XX Lucas A, Michalak M;

XX WPI; 1997-012036/01.

XX Inhibition of restenosis in patients - using calreticulin or a

XX C-domain polypeptide of calreticulin or a variant with the same

XX activity.

XX Disclosure; Fig 1; 48pp; English.

XX The present sequence is calreticulin. It and a C-domain derived peptide

XX (AAW06736) are useful for treating a patient to inhibit restenosis. The

XX calreticulin-type cpds. are administered either parenterally,

XX intravenously or via a catheter and can target areas of vascular damage

XX to inhibit or prevent restenosis.

XX SQ Sequence 401 AA;

Query Match 100.0%; Score 327; DB 18; Length 401;
 Best Local Similarity 100.0%; Pred. No. 1.6e-37;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
 Db 120 cpgptkkvkvifnykgnvlinkdirckddefthlytlivrpdpntyevkidnsqvesgsl 179
 QY 61 E 61
 Db 180 e 180

RESULT 5

ID AAP92276
 AC AAP92276 standard; protein; 417 AA.

XX 23-FEB-1990 (first entry)

XX 60 kD Ro (Ro/SSA) antigen.

XX Sjorens syndrome; systemic lupus erythematosus.

XX Synthetic.

XX WO8909273-A.

XX 05-OCT-1989.

XX 22-MAR-1989; 89WO-US01213.

XX 22-MAR-1988; 88US-0171634.

XX (TEXA) UNIV OF TEXAS SYST.

XX Sontheimer RD, Capra JD, McCauliffe DP;

XX WPI; 1989-309537/42.

XX N-PSDB; AAP92276.

XX DNA sequences encoding antigenic epitope(s) of Ro 60 kD autoantigen

XX used in immunoassays to detect rheumatic disease

XX Disclosure; Fig 2; 88pp; English.

XX Synthetic peptides corresp. to an epitopic core of Ro antigen are

XX expressed recombinantly to detect autoantibodies, for identification

XX of autoimmune diseases. These epitopes are AAs 24-36, 23-36, 188-209,

XX or 241-255. The peptides may be substd. for ribonucleoprotein particle

XX antigens.

XX SQ Sequence 417 AA;

Query Match 100.0%; Score 327; DB 10; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.7e-37;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
 Db 137 cpgptkkvkvifnykgnvlinkdirckddefthlytlivrpdpntyevkidnsqvesgsl 196
 QY 61 E 61
 Db 197 e 197

```

RESULT 6
AAY00927
ID AAY00927 standard; Protein; 417 AA.
XX
AC AAY00927;
XX
DT 28-MAY-1999 (first entry)
XX
DE Calreticulin.
XX
KW Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
KW CUB functionality; inhibitor; complement activation; inflammation;
KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
KW immune complex nephritis; therapy.
XX
DE Homo sapiens.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..17
FT /label= signal_peptide
FT Protein 18
FT /label= mature_protein
XX
PN WO200020577-A1.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US23240.
XX
PR 06-OCT-1998; 98US-0103438.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Tosato G, Pike SE, Yao L;
XX
DR WPI; 2000-303767/26.
XX
DR N-PSDB; AAA09346, AAA09347.
XX
PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
PT useful for suppressing tumor growth
XX
PS Disclosure; Page 79-80; 99pp; English.
XX
CC A novel method of inhibiting endothelial cell growth comprises
CC contacting the cells with calreticulin (or its fragments/variants).
CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
CC tumor growth and/or endothelial cell growth (claimed). The method may be
CC used for inhibiting angiogenesis in a patient. The angiogenesis is
CC associated with a disease other than a tumor that is associated with
CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
CC inflammation, atherosclerosis, excessive wound repair, retinal
CC neovascularization, macular degeneration, corneal graft rejection,
CC contact lens overwear, Crohn's disease, non-immune inflammation,
CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
CC also be used for treating/inhibiting tumor growth especially
CC Kaposi's sarcoma (claimed).
XX
SQ Sequence 417 AA;

Query Match 100.0%; Score 327; DB 21; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.7e-37;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60
Db 137 cpgptkhhvifnykgknlvlnkdrckddethlytlivrpndtyevkidnsqvesgsl 196

QY 61 E 61
Db 197 e 197

RESULT 7
AAY92349
ID AAY92349 standard; Protein; 417 AA.
XX
AC AAY92349;
XX

```


XX AC AAY92354;
 XX 10-AUG-2000 (first entry)
 XX Recombinant human calreticulin residues 121-180.
 DE MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytosolic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200020577-A1.
 XX 13-APR-2000.
 XX 05-OCT-1999; 99WO-US23240.
 XX 06-OCT-1998; 98US-0103438.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Tosato G, Pike SE, Yao L;
 XX WPI; 2000-303767/26.
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth
 XX Claim 4; Page 85; 99pp; English.
 XX A novel method of inhibiting endothelial cell growth comprises
 XX contacting the cells with calreticulin (or its fragments/variants).
 XX Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 XX tumor growth and/or endothelial cell growth (claimed). The method may be
 XX used for inhibiting angiogenesis in a patient. The angiogenesis is
 XX associated with a disease other than a tumor that is associated with
 XX neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 XX trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 XX inflammation, atherosclerosis, excessive wound repair, retinal
 XX neovascularization, macular degeneration, corneal graft rejection,
 XX contact lens overwear, Crohn's disease, non-immune inflammation,
 XX rheumatoid arthritis, systemic lupus erythromatosus, thyroiditis,
 XX Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 XX syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 XX also be used for treating/inhibiting tumor growth especially
 XX Kaposi's sarcoma (claimed).
 XX SQ Sequence 60 AA;
 Query Match 97.2%; Score 318; DB 21; Length 60;
 Best Local Similarity 100.0%; Pred. No. 2.5e-37;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 61
 Db 1 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 60
 RESULT 9
 AAY92355
 ID AAY92355 standard; Protein; 280 AA.
 XX AC AAY92355;
 XX 10-AUG-2000 (first entry)
 XX Recombinant delta-120 calreticulin.
 DE MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytosolic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200020577-A1.
 XX 13-APR-2000.
 XX 05-OCT-1999; 99WO-US23240.
 XX 06-OCT-1998; 98US-0103438.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Tosato G, Pike SE, Yao L;
 XX WPI; 2000-303767/26.
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth
 XX Claim 4; Page 85; 99pp; English.
 XX A novel method of inhibiting endothelial cell growth comprises
 XX contacting the cells with calreticulin (or its fragments/variants).
 XX Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 XX tumor growth and/or endothelial cell growth (claimed). The method may be
 XX used for inhibiting angiogenesis in a patient. The angiogenesis is
 XX associated with a disease other than a tumor that is associated with
 XX neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 XX trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 XX inflammation, atherosclerosis, excessive wound repair, retinal
 XX neovascularization, macular degeneration, corneal graft rejection,
 XX contact lens overwear, Crohn's disease, non-immune inflammation,
 XX rheumatoid arthritis, systemic lupus erythromatosus, thyroiditis,
 XX Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 XX syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 XX also be used for treating/inhibiting tumor growth especially
 XX Kaposi's sarcoma (claimed).
 XX SQ Sequence 60 AA;
 Query Match 97.2%; Score 318; DB 21; Length 60;
 Best Local Similarity 100.0%; Pred. No. 2.5e-37;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 61
 Db 1 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 60
 RESULT 9
 AAY92355
 ID AAY92355 standard; Protein; 280 AA.
 XX AC AAY92355;
 XX 10-AUG-2000 (first entry)
 XX Recombinant delta-120 calreticulin.
 DE MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytosolic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200020577-A1.
 XX 13-APR-2000.
 XX 05-OCT-1999; 99WO-US23240.
 XX 06-OCT-1998; 98US-0103438.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Tosato G, Pike SE, Yao L;
 XX WPI; 2000-303767/26.
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth
 XX Claim 4; Page 86; 99pp; English.
 XX This sequence comprises recombinant human calreticulin (AAY92350)
 XX missing the N-terminal 120 amino acids.
 XX A novel method of inhibiting endothelial cell growth comprises
 XX contacting the cells with calreticulin (or its fragments/variants).
 XX Fragments of calreticulin causes at least 40% inhibition of
 XX angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 XX method may be used for inhibiting angiogenesis in a patient. The
 XX angiogenesis is associated with a disease other than a tumor that is
 XX associated with neovascularization (e.g. diabetic neuropathy, retrolental
 XX fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 XX immune inflammation, atherosclerosis, excessive wound repair, retinal
 XX neovascularization, macular degeneration, corneal graft rejection,
 XX contact lens overwear, Crohn's disease, non-immune inflammation,
 XX rheumatoid arthritis, systemic lupus erythromatosus, thyroiditis,
 XX Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 XX syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 XX also be used for treating/inhibiting tumor growth especially
 XX Kaposi's sarcoma (claimed).
 XX SQ Sequence 280 AA;
 Query Match 97.2%; Score 318; DB 21; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1.8e-36;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 61
 Db 1 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 60
 RESULT 10
 AAW04171
 ID AAW04171 standard; Protein; 403 AA.
 XX AC AAW04171;
 XX 12-DEC-1996 (first entry)
 XX Flea calreticulin PctCal403.
 XX Calreticulin; flea; haematophagous insect; allergic dermatitis;
 KW vaccine; therapy; PctCal403.
 XX Ctenocephalides felis.
 XX

KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytosolic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200020577-A1.
 XX 13-APR-2000.
 XX 05-OCT-1999; 99WO-US23240.
 XX 06-OCT-1998; 98US-0103438.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Tosato G, Pike SE, Yao L;
 XX WPI; 2000-303767/26.
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth
 XX Claim 4; Page 86; 99pp; English.
 XX This sequence comprises recombinant human calreticulin (AAY92350)
 XX missing the N-terminal 120 amino acids.
 XX A novel method of inhibiting endothelial cell growth comprises
 XX contacting the cells with calreticulin (or its fragments/variants).
 XX Fragments of calreticulin causes at least 40% inhibition of
 XX angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 XX method may be used for inhibiting angiogenesis in a patient. The
 XX angiogenesis is associated with a disease other than a tumor that is
 XX associated with neovascularization (e.g. diabetic neuropathy, retrolental
 XX fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 XX immune inflammation, atherosclerosis, excessive wound repair, retinal
 XX neovascularization, macular degeneration, corneal graft rejection,
 XX contact lens overwear, Crohn's disease, non-immune inflammation,
 XX rheumatoid arthritis, systemic lupus erythromatosus, thyroiditis,
 XX Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 XX syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 XX also be used for treating/inhibiting tumor growth especially
 XX Kaposi's sarcoma (claimed).
 XX SQ Sequence 280 AA;
 Query Match 97.2%; Score 318; DB 21; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1.8e-36;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 61
 Db 1 GPGTKKVVHVFNYGKNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 60
 RESULT 10
 AAW04171
 ID AAW04171 standard; Protein; 403 AA.
 XX AC AAW04171;
 XX 12-DEC-1996 (first entry)
 XX Flea calreticulin PctCal403.
 XX Calreticulin; flea; haematophagous insect; allergic dermatitis;
 KW vaccine; therapy; PctCal403.
 XX Ctenocephalides felis.
 XX

```

PN WO9628469-A1.
XX 19-SEP-1996.
XX 08-MAR-1996; 96WO-US03133.
XX 09-MAR-1995; 95US-0401509.
XX (HESK-) HESKA CORP.
XX Rushlow KE, Stiegler GL;
XX WPI; 1996-442861/44.
DR N-PSDB; AAT39516;
DR N-PSDB; AAT39517.
XX Haematophagous insect calreticulin protein - used to reduce insect
PT infestation and desensitise patients to allergic dermatitis
XX Claim 5; Page 68-69; 86pp; English.
XX Flea calreticulin protein PctCall589 (AAW04171) is a calcium-binding
CC protein found in the salivary glands of Ctenocephalides felis
CC Its amino acid sequence was deduced from a cDNA clone (AAT39516)
CC obtg. from a salivary gland cDNA library. Recombinant PctCall589
CC can be produced in host cells transfected with a vector carrying
CC calreticulin nucleic acids. Calreticulin alters the blood feeding
CC behaviour of haematophagous insects and can be administered to an
CC animal to reduce infestation. It reduces calreticulin activity in
CC insects, so reducing the insect burden on an animal. Calreticulin
CC can be used to elicit an immune response, thereby desensitising an
CC animal to allergic dermatitis caused by fleas, mosquitoes or
CC Culicoides.
XX Sequence 403 AA;

Query Match 85.6%; Score 280; DB 17; Length 403;
Best Local Similarity 82.0%; Pred. No. 6.3e-31;
Matches 50; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGPGTKKHVHVIENYKGNVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60
DB 137 cdpgtkxhviisfsgkgnvinkdirckddvthvylvvpdntyevidnekvesgnl 196

QY 61 E 61
DB 197 e 197

RESULT 11
AAY92353
ID AAY92353 standard; Protein; 49 AA.
XX AAY92353;
AC AAY92353;
XX 10-AUG-2000 (first entry)
DT DT
XX Recombinant human calreticulin residues 132-180.
DE DE
XX MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
KW cytosolic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
XX Homo sapiens.
OS Synthetic.
XX WO2000020577-A1.
XX 13-APR-2000.
XX 05-OCT-1999; 99WO-US23240.

XX 06-OCT-1998; 98US-0103438.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Tosato G, Pike SE, Yao L;
XX WPI; 2000-303767/26.
XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
PT useful for suppressing tumor growth
XX Claim 4; Page 82-83; 99pp; English.
XX A novel method of inhibiting endothelial cell growth comprises
CC contacting the cells with calreticulin (or its fragments/variants).
CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
CC tumor growth and/or endothelial cell growth (claimed). The method may be
CC used for inhibiting angiogenesis in a patient. The angiogenesis is
CC associated with a disease other than a tumor that is associated with
CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
CC inflammation, atherosclerosis, excessive wound repair, retinal
CC neovascularization, macular degeneration, corneal graft rejection,
CC contact lens overwear, Crohn's disease, non-immune inflammation,
CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
CC also be used for treating/inhibiting tumor growth especially
CC Kaposi's sarcoma (claimed).
XX Sequence 49 AA;

Query Match 78.9%; Score 258; DB 21; Length 49;
Best Local Similarity 100.0%; Pred. No. 5.1e-29;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 NYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 61
DB 1 nykgnvlinkdirckddefthlytlivrpdttyevkidnsqvesgsl 49

RESULT 12
AAR12312
ID AAR12312 standard; Protein; 336 AA.
XX AAR12312;
AC AAR12312;
XX 29-AUG-1991 (first entry)
DT DT
XX Partial sequence of Onchocera volvulus 42 kD antigen.
DE DE
XX River blindness; onchocerciasis; vaccine; antigen; parasite.
KW Onchocerca volvulus.
XX Onchocerca volvulus.
XX Key Location/Qualifiers
FT Region 160..165
FT /label= repeat unit
FT /note= "hydrophilic"
FT Region 177..183
FT /label= repeat unit
FT /note= "hydrophilic"
FT Region 195..201
FT /label= repeat unit
FT /note= "hydrophilic"
XX US5021342-A.
XX 04-JUN-1991.
XX 30-JUN-1988; 88US-0214264.

```

```

XX 30-JUN-1988; 88US-0214264.
XX PA (UYHO-) UNIV HOSPITALS CLEV.
XX PI
XX XX Greene BM, Unnasch TR;
XX DR WPI; 1991-185179/25.
XX DR N-PSDB; AAQ11987.
XX PT DNA encoding Onchocerca volvulus antigen - used to express
XX PT recombinant antigen for vaccine against onchocerciasis or river
XX PT blindness.
XX PS Disclosure; Fig 7; 20pp; English.
XX CC The sequence was deduced from a cDNA clone lambda RAL-1 prep'd.
XX CC from RNA isolated from nodules excised from patients infected
XX CC with O. volvulus. The N-terminal is incomplete, however Abs
XX CC which specifically bind to protein prep'd. from induced cultures
XX CC of lambda RAL-1 lysogens recognise a single polypeptide of mol.
XX CC wt. 42,000 in extracts of adult worms. Analysis deduced sequence
XX CC suggests that it encodes a protein of mol. wt. 39,130. If the
XX CC antigen is not subject to post-translational processing this
XX CC suggests that most of the coding sequence is present. The three
XX CC repeats are highly hydrophilic regions likely to be exposed on
XX CC the surface of the antigen and highly immunogenic. Recombinant
XX CC antigen expressed by the clone can be used stimulate T-cells of
XX CC individuals infected by the parasite to proliferate and may be
XX CC used as the basis for a vaccine against Onchocerciasis or river
XX CC blindness.
XX SQ Sequence 336 AA;

Query Match 76.8%; Score 251; DB 12; Length 336;
Best Local Similarity 75.4%; Pred. No. 5.9e-27;
Matches 46; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGPQTKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSL 60
DB 83 cgpqtkkvhvifnykdrnhmkdkirckdvfthlytlivnsdntyevqidgkaesgl 142

QY 61 E 61
DB 143 e 143

RESULT 13
AAY00924
ID AAY00924 standard; Protein; 122 AA.
XX AC AAY00924;
XX DT 28-MAY-1999 (first entry)
XX DE Human cClqR binding domain protein sequence.
XX KW Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
XX KW CUB functionality; inhibitor; complement activation; inflammation;
XX KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
XX KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
XX KW Immune complex nephritis; therapy.
XX OS Homo sapiens.
XX PN WO9907406-A1.
XX PD 18-FEB-1999.
XX PF 12-AUG-1998; 98WO-GB02430.
XX PR 12-AUG-1997; 97GB-0016998.

XX PA (UYLE-) UNIV LEICESTER.
XX PI
XX XX Schwaeble W;
XX DR WPI; 1999-180404/15.
XX DR N-PSDB; AAX27251.
XX PT Use of a cClqR binding domain - to modulate complement ubiquitin
XX PT (CUB) functionality.
XX PS Claim 9; Page 23; 31pp; English.
XX CC This sequence is a Clq and collectin receptor (cClqR) binding
XX CC domain. The invention relates to the use of a cClqR binding domain in a
XX CC medicament to effect complement ubiquitin (CUB) functionality, and an
XX CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
XX CC functionality. The cClqR binding domain, or its inhibitor, can be used to
XX CC treat a human or animal body. Particularly an inhibitor is used to treat
XX CC complement activation involved in the initiation and maintenance of
XX CC inflammation, for example in myocardial infarction, brain ischaemia
XX CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
XX CC erythematosus, burns, immune complex nephritis, and to treat amyloid
XX CC plaques in Alzheimer's disease. The use of cClqR binding domain or
XX CC inhibitor enables the CUB domain functionality to be modulated using a
XX CC low molecular weight molecule.
XX SQ Sequence 122 AA;

Query Match 58.1%; Score 190; DB 20; Length 122;
Best Local Similarity 100.0%; Pred. No. 5.8e-19;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 RCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 61
DB 1-rckddefthlytlivrpndtyevkldnsqvsgsle 36

RESULT 14
AAY00926
ID AAY00926 standard; Protein; 122 AA.
XX AC AAY00926;
XX DT 28-MAY-1999 (first entry)
XX DE Rat cClqR binding domain protein sequence.
XX KW Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
XX KW CUB functionality; inhibitor; complement activation; inflammation;
XX KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
XX KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
XX KW Immune complex nephritis; therapy.
XX OS Rattus norvegicus.
XX PN WO9907406-A1.
XX PD 18-FEB-1999.
XX PF 12-AUG-1998; 98WO-GB02430.
XX PR 12-AUG-1997; 97GB-0016998.

XX PA (UYLE-) UNIV LEICESTER.
XX PI
XX XX Schwaeble W;
XX DR WPI; 1999-180404/15.
XX DR N-PSDB; AAX27253.
XX PT Use of a cClqR binding domain - to modulate complement ubiquitin
XX PT
```

PT (CUB) functionality.
XX
PS Claim 9; Page 24-25; 3lpp; English.
XX
CC This sequence is a C1q and collectin receptor (cClqR) binding
CC domain. The invention relates to the use of a cClqR binding domain in a
CC medicament to effect complement ubiquitin (CUB) functionality, and an
CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
CC functionality. The cClqR binding domain, or its inhibitor, can be used to
CC treat a human or animal body. Particularly, an inhibitor is used to treat
CC complement activation involved in the initiation and maintenance of
CC inflammation, for example in myocardial infarction, brain ischaemia
CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
CC erythematosus, burns, immune complex nephritis, and to treat amyloid
CC plaques in Alzheimer's disease. The use of cClqR binding domain or
CC inhibitor enables the CUB domain functionality to be modulated using a
CC low molecular weight molecule.
XX
SQ Sequence 122 AA;

Query Match 58.1%; Score 190; DB 20; Length 122;
Best Local Similarity 100.0%; Pred. No. 5.8e-19;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 RCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSL 61
DB 1 rckddefthlytlivrpdntyevkidnsqvesgsl 36

RESULT 15
AAB66341
ID AAB66341 standard; Protein; 415 AA.
XX
AC AAB66341;
XX
DT 05-APR-2001 (first entry)
XX
DE Castor bean calreticulin SEQ ID NO: 2.
XX
KW Castor bean; calreticulin; calcium binding; er; endoplasmic reticulum;
KW calnexin; promoter; resistance.
XX
OS Ricinus communis.
XX
PN US6171864-B1.
XX
PD 09-JAN-2001.
XX
PF 05-JUL-1996; 96US-0675816.
XX
PR 05-JUL-1996; 96US-0675816.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Coughlan SJ, Winfrey RJ;
XX
DR WPI: 2001-122335/13.
DR N-PSDB; AAF29741, AAF29742.
XX
PT New nucleic acid molecules encoding a calcium binding chaperone protein
PT in endoplasmic reticulum, calreticulin, and calreticulin promoter
PT sequences, useful for producing foreign gene products in plant cells -
XX
PS Claim 1; Fig 3; 45pp; English.
XX
CC The present invention provides the protein and coding sequences for the
CC castor bean calreticulin protein and the calreticulin promoter sequence.
CC In addition, the castor bean calnexin protein, coding sequence and
CC promoter are also described. Calreticulin and calnexin are calcium
CC binding proteins found in the endoplasmic reticulum (er). The
CC calreticulin promoter sequence can be used in vectors to promote the
CC expression of foreign genes, particularly resistance genes, in plant

CC cells.
XX
SQ Sequence 415 AA;
Query Match 56.9%; Score 186; DB 22; Length 415;
Best Local Similarity 53.3%; Pred. No. 1e-17;
Matches 32; Conservative 11; Mismatches 17; Indels 0; Gaps 0;
QY 1 CGPCTKKVHVIFNYKGNVLIINKDIRKDEFFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
DB 138 cgystkkvhailnyndnlnhlikkevpcetdqlthvylvirpdntysilidnvektgsl 197
Search completed: October 21, 2001, 03:05:47
Job time: 342 sec

THIS PAGE BLANK (USPTO)


```

; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/675,816
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)-622-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-675-816-6

Query Match          25.2%; Score 82.5; DB 4; Length 542;
Best Local Similarity 31.7%; Pred.No.0.004;
Matches 19; Conservative 10; Mismatches 24; Indels 7; Gaps
QY      1 CGPGTKKHVIFNVKG--KNVLINKDIR----CKDDEFTHLYTLIVRPDMTYEVKIDNSQ 54
           ||| | | | : - : - : - : - : - : - : - : - : - : - : - : - : - : - : - : - :
Db       147 CG-ATNKVHPILHKPKSEYTEHLKYPPSPDKLTHVTAILKPDELRLVDGEE 205


RESULT            4
US-08-906-865-3
Sequence 3, Application US/08906865
Patent No. 6040168
GENERAL INFORMATION:
APPLICANT: Greengard, Paul
APPLICANT: Porton, Barbara
APPLICANT: Kao, Hung-Teh
TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave., Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,865
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq.; David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 582 amino acids
```


RESULT 8
 US-07-994-423-2
 ; Sequence 2, Application US/07994423
 ; Patent No. 5432080
 ; GENERAL INFORMATION:
 ; APPLICANT: HAMMOND, Geoffrey L.
 ; APPLICANT: AVAKUMOV, George V.
 ; TITLE OF INVENTION: Variants of Corticosteroid Binding
 ; TITLE OF INVENTION: Globulin
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/994.423
 ; FILING DATE: 16-DEC-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16777/197/ALLE
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 405 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-07-994-423-2

```

Query Match      16.5%;   Score 54;   DB 1;   Length 405;
Best Local Similarity 36.7%;   Pred. NO. 21;
Matches 11; Conservative 6; Mismatches 13; Indels

QY    32 FTHLYTLIVRPONTVEVKIDNSQVESGSL E 61
      : ||| : | : | : | : | : | : ||||
Db     109 FQHLHQLEAKSDTSLEMTGMNALFLDGSLE 138

RESULT 10
US-08-190-802A-54
; Sequence 54, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Moehly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof

```

```
;
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: MUS MUSCULUS PROTEIN, Fig. 37
; US-08-190-802A-54

Query Match 16.5%; Score 54; DB 1; Length 816;
Best Local Similarity 29.8%; Pred. No. 51;
Matches 14; Conservative 8; Mismatches 15; Indels 10; Gaps 2;

QY 10 VFN-----YKGNVLINKDIRCKDDEFTHLVTLVLRPNTYEVKI 50
   :||:  :  :||:  :  :||:  :  :||:  :  :||:  :
Db 536 VFNIDLRQAHFASKLVIKD----GDKKVLTVFVNPANVQFAV 578

RESULT 11
US-08-701-191A-23
; Sequence 23, Application US/08701191A
; Patent No. 5942428
; GENERAL INFORMATION:
; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
; APPLICANT: and Stevan R. Hubbard
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
; OF NON-INSULIN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
```

```
;
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEEX: 67-3510
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-701-191A-23

Query Match 16.4%; Score 53.5; DB 2; Length 279;
Best Local Similarity 36.4%; Pred. No. 15;
Matches 12; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 17 KNVLINKDIRCKDDEFTHLVTLVLRPNTYEVK 49
   :||:  :  :||:  :  :||:  :  :||:  :
Db 143 RNILVNQLCKVSDF-GLRLDLDGDTYETQ 174

RESULT 12
US-08-292-549-6
; Sequence 6, Application US/08292549
; Patent No. 5464938
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,549
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/963,330
; FILING DATE: 10/19/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2602-A
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-292-549-6
```

;; TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC

Search completed: October 21, 2001, 03:06:41
Job time: 331 sec

THIS PAGE BLANK (USPTO)

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	327	100.0	416	1	S06763	calreticulin precu
2	327	100.0	416	2	JH0819	calreticulin precu
3	327	100.0	417	1	A37047	calreticulin precu
4	327	100.0	418	1	A34154	calreticulin precu
5	322	98.5	400	2	A34376	calreticulin, brai
6	322	98.5	421	1	A36799	calreticulin precu
7	290	88.7	405	1	JH0795	calreticulin precu
8	287	87.8	384	2	S29130	calreticulin (clon
9	287	87.8	411	2	S29129	calreticulin precu
10	278	85.0	406	2	A36637	calreticulin homol
11	271	82.9	419	2	S71343	calreticulin precu
12	251	76.8	336	2	A32507	41K larval antigen
13	236	72.2	393	1	A48573	calreticulin autoa
14	219.5	67.1	395	2	S25851	calreticulin precu
15	193	59.0	416	2	T14534	calreticulin - bee
16	188	57.5	421	2	S58170	calreticulin precu
17	186	56.9	415	2	T10172	calreticulin - cas
18	185	56.6	412	2	T05703	calreticulin - bar
19	185	56.6	415	2	T05705	calreticulin - bar
20	181	55.4	425	2	C36605	calreticulin (Ctrl
21	178	54.4	389	2	T03691	calreticulin - com
22	178	54.4	416	2	T16968	calreticulin call
23	178	54.4	442	2	H86224	calreticulin call
24	177	54.1	422	2	T07841	probable calreticu
25	117.5	35.9	592	2	I53260	calnexin - human
26	117.5	35.9	592	2	A46673	calnexin precursor
27	116.5	35.6	591	2	B34354	calnexin precursor
28	116.5	35.6	591	2	C54354	calnexin precursor
29	115.5	35.3	593	1	A37273	calnexin precursor

R.Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; M
 Biochem. J. 271, 473-480, 1990
 A:Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-stor
 A:Reference number: S13045; MUID:91054414
 A:Accession: S13045
 A:Molecule type: protein
 A:Residues: 18-29 <TRE>
 C:Superfamily: calreticulin
 C:Keywords: calcium binding; glycoprotein
 F:1-17/Domain: signal sequence status predicted <SIG>
 F:18-416/Product: calreticulin status experimental <MAT>
 F:204-212/Region: nuclear location signal
 F:413-416/Region: endoplasmic reticulum retention signal
 F:344/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 327; DB 2; Length 416;
 Best Local Similarity 100.0%; Pred. No. 1.le-30;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKVVHVIENYKGNVLIINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSL 60
 Db 137 CGPGTKKVVHVIENYKGNVLIINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSL 196
 QY 61 E 61
 Db 197 E 197

RESULT 3
 A37047
 calreticulin precursor - human
 N:Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding pr
 C:Species: Homo sapiens (man)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Feb-2000
 C:Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075
 R:McCaulliffe, D.P.; Yang, Y.S.; Wilson, J.; Sontheimer, R.D.; Capra, J.D.
 J. Biol. Chem. 267, 2557-2562, 1992
 A:Title: The 5'-flanking region of the human calreticulin gene shares homology with t
 A:Reference number: A42330; MUID:92129342
 A:Accession: A42330
 A:Molecule type: DNA
 A:Residues: 1-417 <MC2>
 A:Note: sequence extracted from NCBI backbone (NCBIN:78537, NCBIP:78536)
 R:McCaulliffe, D.P.; Lux, P.A.; Lieu, T.S.; Sanz, J.; Hanke, J.; Newkirk, M.M.; Bachin
 J. Clin. Invest. 85, 1379-1391, 1990
 A:Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/
 A:Reference number: A37047; MUID:90237213
 A:Accession: A37047
 A:Molecule type: mRNA
 A:Residues: 1-417 <MCC>
 A:Cross-references: GB:M32294; NID:g337486; PIDN:AAA36582.1; PID:g337487
 A:Note: the authors translated the codon GFA for residue 349 as tyr
 R:Roakeach, L.A.; Haseilby, J.A.; Meilof, J.F.; Smeenk, R.J.; Unnasch, T.R.; Greene, B.
 J. Immunol. 147, 3031-3039, 1991
 A:Title: Characterization of the autoantigen calreticulin.
 A:Reference number: A46452; MUID:92013129
 A:Accession: A46452
 A:Molecule type: mRNA
 A:Residues: 1-417 <ROK>
 A:Cross-references: GB:M84739; NID:g179881; PIDN:AAA51916.1; PID:g179882
 A:Note: sequence extracted from NCBI backbone (NCBIN:60749, NCBIP:60750)
 R:Lieu, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.
 J. Clin. Invest. 82, 96-101, 1988
 A:Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence
 A:Reference number: A28812; MUID:88273610
 A:Accession: A28812
 A:Molecule type: protein
 A:Residues: 18-41 <LIE>
 A:Note: 18-Ala was also found
 R:Dupuis, M.; Schaerer, E.; Krause, K.H.; Tschopp, J.
 J. Exp. Med. 177, 1-7, 1993
 A:Title: The calcium-binding protein calreticulin is a major constituent of lytic gra

RESULT 2
 JH0819
 calreticulin precursor - rat
 N:Alternate names: calcium-binding protein 3
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
 C:Accession: JH0819; A49176; S45036; S04867; S39372; A34473; S13045
 R:Nakamura, M.; Moriya, M.; Baba, T.; Michikawa, Y.; Yamanobe, T.; Arai, K.; Okinaga, S.
 Exp. Cell Res. 205, 101-110, 1993
 A:Title: An endoplasmic reticulum protein, calreticulin, is transported into the acrosom
 A:Reference number: A49176; MUID:93302172
 A:Accession: JH0819
 A:Molecule type: mRNA
 A:Residues: 1-416 <NAK>
 A:Cross-references: GB:D78308; NID:g1089798; PIDN:BA11345.1; PID:g1845572
 A:Accession: A49176
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-416 <NA2>
 A:Cross-references: GB:D78308; NID:g1089798; PIDN:BA11345.1; PID:g1845572
 A:Experimental source: Sprague-Dawley, spermatogenic cells
 A:Note: sequence extracted from NCBI backbone (NCBIN:127639, NCBIP:127643)
 R:Murthy, K.K.; Banville, D.; Srikant, C.B.; Carrier, F.; Holmes, C.; Bell, A.; Patel, Y
 Nucleic Acids Res. 18, 4933, 1990
 A:Title: Structural homology between the rat calreticulin gene product and the Onchocerc
 A:Reference number: S11205; MUID:90370496
 A:Accession: S11205
 A:Molecule type: mRNA
 A:Residues: 1-416 <NUR>
 A:Cross-references: EMBL:X53363; NID:g55854; PIDN:CAA37446.1; PID:g55855
 R:Nakamura, M.; Michikawa, Y.; Baba, T.; Okinaga, S.; Arai, K.
 Biochem. Biophys. Res. Commun. 186, 668-673, 1992
 A:Title: Calreticulin is present in the acrosome of spermatids of rat testis.
 A:Reference number: PC1109; MUID:92360010
 A:Accession: PC1109
 A:Molecule type: protein
 A:Residues: 18-32 <NAK2>
 A:Experimental source: testis, strain Sprague-Dawley
 R:Soennichsen, B.; Fuellekrug, J.; van Nguyen, P.; Diekmann, W.; Robinson, D.G.; Mieskes
 submitted to the EMBL Data Library, May 1994
 A:Description: Retention and retrieval: both mechanisms cooperate to maintain calreticul
 A:Reference number: S45036
 A:Accession: S45036
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-416 <SOE>
 A:Cross-references: EMBL:X79327; NID:g488840; PIDN:CAA55890.1; PID:g488841
 R:Lone, Y.C.; Bailly, A.; Latruffe, N.
 submitted to the EMBL Data Library, December 1988
 A:Reference number: S04867
 A:Accession: S04867
 A:Molecule type: mRNA
 A:Residues: 'R', 270-358, 'AAG' <ION>
 A:Cross-references: EMBL:X13702; NID:g56055; PIDN:CAA31987.1; PID:g930260
 A:Note: the authors designated the protein as D-beta-hydroxybutyrate dehydrogenase
 R:Yokoi, T.; Nagayama, S.; Kajiwara, R.; Kawaguchi, Y.; Horiuchi, R.; Kamataki, T.
 Biochim. Biophys. Acta 1158, 339-344, 1993
 A:Title: Identification of protein disulfide isomerase and calreticulin as autoimmune an
 A:Reference number: S39371; MUID:94072621
 A:Accession: S39372
 A:Molecule type: protein
 A:Residues: 18-23 'X', 25-32 <YOK>
 R:Van, P.N.; Peter, F.; Soelling, H.D.
 J. Biol. Chem. 264, 17494-17501, 1989
 A:Title: Four intracisternal calcium-binding glycoproteins from rat liver microsomes wit
 itive calcium sequestering rat liver vesicles.
 A:Reference number: A34473; MUID:90008920
 A:Accession: A34473
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 18-36 <VAN>

A:Reference number: PH1525; MUID:931115648

A:Accession: PH1525

A:Molecule type: protein

A:Residues: 18-27 <DUP>

A:Experimental source: LAK cell

R:RoJiani, M.V.; Finlay, B.B.; Gray, V.; Dedhar, S.

Biochemistry 30, 9859-9866, 1991

A:Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (calreticulin precursor, skeletal muscle - rabbit)

A:Reference number: A40346; MUID:92002034

A:Accession: A40346

A:Molecule type: protein

A:Residues: 18-34, 'R' <ROJ>

R:Krause, J.H.; Simmerman, H.K.B.; Jones, L.R.; Campbell, K.P.

Biochem. J. 270, 545-548, 1990

A:Title: Sequence similarity of calreticulin with a Ca(2+)-binding protein that co-purifies with calreticulin

A:Reference number: S11475; MUID:90380058

A:Accession: S11475

A:Molecule type: protein

A:Residues: 18-32 <KRA>

R:Lamerdin, J.; McCreedy, P.; Stillwagen, S.; Ramirez, M.; Carrano, A.

Submitted to the EMBL Data Library, November 1996

A:Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region

A:Reference number: 222906

A:Accession: T45075

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-417 <LAW>

A:Cross-references: EMBL:AD000092; PIDN:AAB511176.1

A:Experimental source: cell line 5HL2-B; fibroblast

C:Comment: Autoantibodies specific for this protein are found in Sjogren's syndrome and

C:Genetics:

A:Gene: GDB:CALR

A:Cross-references: GDB:125179; OMIM:109091

A:Map position: 19p13.3-19p13.2

A:Introns: 31/1; 65/1; 133/1; 164/3; 234/3; 272/3; 320/3; 351/3

A:Note: CRPC

C:Superfamily: calreticulin

C:Keywords: calcium binding; integrin binding

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-417/Product: calreticulin #status predicted <MAT>

F:414-417/Region: endoplasmic reticulum retention signal

Query Match

Best Local Similarity 100.0%; Score 327; DB 1; Length 417;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKVVHVFYKGNVNLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60

Db 137 CGPGTKKVVHVFYKGNVNLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 196

QY 61 E 61

Db 197 E 197

RESULT 4

A34154

calreticulin precursor, skeletal muscle - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A34154; S13047

R:Fliegel, L.; Burns, K.; MacLennan, D.H.; Reithmeier, R.A.F.; Michalak, M.

J. Biol. Chem. 264, 21522-21528, 1989

A:Title: Molecular cloning of the high affinity calcium-binding protein (calreticulin) c

A:Reference number: A34154; MUID:90094320

A:Accession: A34154

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-418 <FLI>

A:Cross-references: NID:g164858; PIDN:AAA31188.1; PID:g164859

R:Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld

Biochem. J. 271, 473-480, 1990

A:Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-stor

A:Reference number: S13045; MUID:91054414

A:Accession: S13047

A:Molecule type: protein

A:Residues: 19-32 <TRE>

C:Superfamily: calreticulin

C:Keywords: skeletal muscle

F:1-17/Domain: signal sequence #status predicted <SIG>

F:415-418/Region: endoplasmic reticulum retention signal

Query Match

Best Local Similarity 100.0%; Score 327; DB 1; Length 418;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1'CGPGTKKVVHVFYKGNVNLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60

Db 137 CGPGTKKVVHVFYKGNVNLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 196

QY 61 E 61

Db 197 E 197

RESULT 5

S43376

calreticulin, brain isoform 1 - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 20-Oct-1994 #sequence_revision 23-Mar-1995 #text_change 07-May-1999

C:Accession: S43376; S36801

R:Matshoka, K.; Seta, K.; Yamakawa, Y.; Okuyama, T.; Shinoda, T.; Isobe, T.

Biochem. J. 298, 435-442, 1994

A:Title: Covalent structure of bovine brain calreticulin.

A:Reference number: S43376; MUID:94183174

A:Accession: S43376

A:Molecule type: protein

A:Residues: 1-400 <MAT>

A:Experimental source: brain

R:Liu, N.; Fine, R.E.; Johnson, R.J.

Biochim. Biophys. Acta 1202, 70-76, 1993

A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticuli

A:Reference number: S36799; MUID:93385184

A:Accession: S36801

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 45-63, 'E', '65-83 <LIU>

A:Superfamily: calreticulin

C:Keywords: calcium binding; glycoprotein

F:397-400/Region: endoplasmic reticulum retention signal

F:120-146/Disulfide bonds: #status experimental

F:162/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match

Best Local Similarity 98.5%; Score 322; DB 2; Length 400;

Matches 60; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKVVHVFYKGNVNLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60

Db 120 CGPGTKKVVHVFYKGNVNLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 179

QY 61 E 61

Db 180 E 180

RESULT 6

S36799

calreticulin precursor, brain isoform 2 - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Dec-1993 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999

C:Accession: S36799; S36800

R;Liu, N.; Fine, R.E.; Johnson, R.J.
 Biochim. Biophys. Acta 1202, 70-76, 1993
 A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin.
 A:Reference number: S36799; MUID:93385184
 A:Accession: S36799
 A:Molecule type: mRNA
 A:Residues: 1-421 <LIU>
 A:Cross-references: GB:L13462; NID:9348693; PIDN:AAC37307.1; PID:9348694
 A:Experimental source: brain, clone 9.4
 A:Accession: S36800
 A:Molecule type: protein
 A:Residues: 35-45 <LI2>
 C:Superfamily: calreticulin
 C:Keywords: calcium binding; glycoprotein
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-421/Product: calreticulin, brain isoform 2 #status predicted <MAT>
 F:418-421/Region: endoplasmic reticulum retention signal
 F:141-167/Disulfide bonds: #status predicted
 F:183/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.5%; Score 322; DB 2; Length 421;
 Best Local Similarity 98.4%; Pred. No. 4.2e-30;
 Matches 60; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVLLINKDKDDEFTHLYTLIVRPNTYEVKIDNSQVSGSL 60
 |||||
 Db 141 CGPGTKKHVIFNYKGNVLLINKDKDDEFTHLYTLIVRPNTYEVKIDNSQVSGSL 200

QY 61 E 61
 Db 201 E 201

RESULT 7
 JH0795
 calreticulin precursor - California sea hare
 N:Alternate names: protein 407
 C:Species: Aplysia californica (California sea hare)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: JH0795; B31409; F60977
 R:Kennedy, T.E.; Kuhl, D.; Barzilai, A.; Sweatt, J.D.; Kandel, E.R.
 Neuron 9, 1013-1024, 1992
 A:Title: Long-term sensitization training in aplysia leads to an increase in calreticulin
 A:Reference number: JH0795; MUID:93098937
 A:Accession: JH0795
 A:Molecule type: mRNA
 A:Residues: 1-405 <KEN>
 A:Cross-references: GB:S51239; NID:9262053; PIDN:AAB24569.1; PID:9262054
 A:Experimental source: abdominal ganglion and antral nervous system
 R:Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.; Sweatt, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 85, 7008-7012, 1988
 A:Title: Sequencing of proteins from two-dimensional gels by using in situ digestion and
 tion in Aplysia
 A:Reference number: A94207; MUID:88320566
 A:Accession: B31409
 A:Molecule type: protein
 A:Residues: 'X', 17-28, 'X', 30-31 <KE2>
 R:Sweatt, J.D.; Kennedy, T.E.; Wager-Smith, K.; Gawinowicz, M.A.; Barzilai, A.; Karl, K.
 Electrophoresis 10, 152-157, 1989
 A:Title: Development of a database of amino acid sequences for proteins identified and i
 A:Reference number: A60977; MUID:89276264
 A:Accession: F60977
 A:Molecule type: protein
 A:Residues: 'X', 17-28, 'X', 30-31 <SWE>
 C:Superfamily: calreticulin
 C:Keywords: calcium binding; endoplasmic reticulum
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-405/Product: calreticulin #status experimental <MAT>
 F:402-405/Region: endoplasmic reticulum retention signal

Query Match 88.7%; Score 290; DB 1; Length 405;

Best Local Similarity 85.2%; Pred. No. 2.3e-26;
 Matches 52; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVLLINKDKDDEFTHLYTLIVRPNTYEVKIDNSQVSGSL 60
 |||||
 Db 133 CGPGTKKHVIFNYKGNVLLINKDKDDEFTHLYTLIVRPNTYEVKIDNSQVSGSL 192

QY 61 E 61
 Db 193 E 193

RESULT 8
 S29129
 calreticulin (clone 8) - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
 C:Accession: S29130; T01068
 R:Treves, S.; Zorzato, F.; Pozzan, T.
 Biochem. J. 287, 579-581, 1992
 A:Title: Identification of calreticulin isoforms in the central nervous system.
 A:Reference number: S29129; MUID:93074997
 A:Accession: S29130
 A:Molecule type: mRNA
 A:Residues: 1-384 <TRE>
 A:Cross-references: EMBL:X67598
 A:Accession: T01068
 A>Status: translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-339, 'XTGR' <TRM>
 A:Cross-references: EMBL:X67598; NID:964610; PIDN:CAA47867.1; PID:964611
 A:Experimental source: CNS
 C:Superfamily: calreticulin
 C:Keywords: glycoprotein
 F:381-384/Region: endoplasmic reticulum retention signal
 F:316/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.8%; Score 287; DB 2; Length 384;
 Best Local Similarity 88.5%; Pred. No. 4.8e-26;
 Matches 54; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVLLINKDKDDEFTHLYTLIVRPNTYEVKIDNSQVSGSL 60
 |||||
 Db 109 CGPGTKKHVIFNYKGNVLLINKDKDDEFTHLYTLIVRPNTYEVKIDNSQVSGSL 168

QY 61 E 61
 Db 169 E 169

RESULT 9
 S29129
 calreticulin precursor (clone 3) - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
 C:Accession: S29129
 R:Treves, S.; Zorzato, F.; Pozzan, T.
 Biochem. J. 287, 579-581, 1992
 A:Title: Identification of calreticulin isoforms in the central nervous system.
 A:Reference number: S29129; MUID:93074997
 A:Accession: S29129
 A:Molecule type: mRNA
 A:Residues: 1-411 <TRE>
 A:Cross-references: EMBL:X67597; NID:964608; PIDN:CAA47866.1; PID:964609
 C:Superfamily: calreticulin
 C:Keywords: glycoprotein
 F:1-12/Domain: signal sequence (fragment) #status predicted <SIG>
 F:13-411/Product: calreticulin #status predicted <MAT>
 F:408-411/Region: endoplasmic reticulum retention signal
 F:339/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.8%; Score 287; DB 2; Length 411;
Best Local Similarity 88.5%; Pred. No. 5.2e-26;
Matches 54; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 132 CGPTTKKVHVIFYQQKNKLQINKDIRCKDDSFTHLYTLIVRPDNTYEVKIDNSQVESGSL 191

QY 61 E 61
↓
Db 192 E 192

RESULT 10
A56637
calreticulin homolog precursor - fruit fly (*Drosophila melanogaster*)
N:Alternate names: Ro/SS-A autoantigen/calreticulin homolog
C:Species: *Drosophila melanogaster*
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 13-Aug-1999
C:Accession: A56637; A37158
R:Smith, M.J.
DNA Seq. 3, 247-250, 1992
A:Title: Nucleotide sequence of a *Drosophila melanogaster* gene encoding a calreticulin homolog
A:Reference number: A56637; MUID:93208374
A:Accession: A56637
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-406 <SMI>
A:Cross-references: GB:X64461; NID:g7685; PIDN:CAA45791.1; PID:g7686
A>Note: sequence extracted from NCBI backbone (NCBIN:128274, NCBIPI:128275)
R:McCaulliffe, D.P.; Zappi, E.; Lieu, T.S.; Michalak, M.; Sontheimer, R.D.; Capra, J.D.
J. Clin. Invest. 86, 332-335, 1990
A:Title: A human Ro/SS-A autoantigen is the homologue of calreticulin and is highly homologous to the mouse Ro/SS-A autoantigen
A:Reference number: A37158; MUID:90307981
A:Accession: A37158
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 91-105,'A',107,109-124;182-183,'L',185-220 <MCC>
C:Genetics:
A:Gene: FlyBase:Crc
A:Cross-references: FlyBase:FBgn0005585
A:Introns: 65/1; 222/3
C:Superfamily: calreticulin
C:Keywords: calcium binding; endoplasmic reticulum
F:1-17/Domain: signal sequence #status predicted <SIG>
F:403-406/Region: endoplasmic reticulum retention signal

Query Match 85.0%; Score 278; DB 2; Length 406;
Best Local Similarity 83.6%; Pred. No. 5.8e-25;
Matches 51; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 CGPCTKKVHVIFYSYKGNHLISKDIRCKDDVYTFHYTLIVRPDNTYEVYLIDNKVESGNL 196

QY 61 E 61
↓
Db 197 E 197

RESULT 11
S71343
calreticulin precursor - Korean frog
C:Species: *Rana rugosa* (Korean frog)
C>Date: 29-Jan-1998 #sequence_revision 13-Feb-1998 #text_change 20-Jun-2000
C:Accession: S71343
R:Yamamoto, S.; Nakamura, M.
FEBS Lett. 387, 27-32, 1996
A:Title: Calnexin: its molecular cloning and expression in the liver of the frog, *Rana rugosa*
A:Reference number: S71342; MUID:96234004
A:Accession: S71343
A>Status: nucleic acid sequence not shown

```

Query Match      72.2%; Score 236; DB 1; Length 393;
Best Local Similarity 73.8%; Pred. No. 4.7e-20;
Matches 45; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGPGTKKVVHVFYFNKGNVINKDKDDETHLYTLVLRPDNTYEVKIDNSQVESGSL 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 CGMATKKVHVIFNYKGNHILKRIKPEKDDKTLKTLVNPNNKYELVDNAKVEGSL 194

QY 61 E 61
Db 195 E 195

RESULT 14
S25851
calreticulin precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S25851; MUID:92329978
R:Smith, M.J.
DNA Seq. 2, 235-240, 1992
A:Title: A C. elegans gene encodes a protein homologous to mammalian calreticulin.
A:Reference number: S25851; MUID:92329978
A:Accession: S25851
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <SMI>
A:Cross-references: EMBL:X59589; NID:g5693; PIDN:CAA42159.1; PID:g6694
R:Bauer, C.; Courtney, L.; Lariant, Y.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid Y38A10A.
A:Reference number: Z21453
A:Accession: T33996
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-395 <BAU>
A:Cross-references: EMBL:AF125963; PIDN:AD14746.1; GSPDB:GN00023; CESP:Y38A10A.5
A:Experimental source: strain Bristol N2; clone Y38A10A
C:Genetics:
A:Gene: CESP:Y38A10A.5
A:Map position: 5
A:Introns: 107/3; 315/3
C:Superfamily: calreticulin
F:1-15/Domain: signal sequence #status predicted <SIG>
F:392-395/Region: endoplasmic reticulum retention signal

Query Match      67.1%; Score 219.5; DB 2; Length 395;
Best Local Similarity 67.2%; Pred. No. 4.1e-18;
Matches 41; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 1 CGPGTKKVVHVFYFNKGNVINKDKDDETHLYTLVLRPDNTYEVKIDNSQVESGSL 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 CGP-TRRVHILNYKGNLKRITCKSDKTLKTLVLRPDNTYEVKIDNSQVESGSL 191

QY 61 E 61
Db 192 E 192

RESULT 15
T14554
calreticulin - beet
C:Species: Beta vulgaris (beet)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
C:Accession: T14554
R:Viereck, R.
submitted to the EMBL Data Library, October 1997
A:Description: Nucleotide sequence from sugar beet calreticulin.
A:Reference number: Z18137
A:Accession: T14554

```

```

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-416 <VIE>
A:Cross-references: EMBL:AJ002057
A:Experimental source: strain diploide Inzuchtlinie KWS; leaf
C:Superfamily: calreticulin
C:Keywords: calcium binding

Query Match      59.0%; Score 193; DB 2; Length 416;
Best Local Similarity 56.7%; Pred. No. 5.6e-15;
Matches 34; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 CGPGTKKVVHVFYFNKGNVINKDKDDETHLYTLVLRPDNTYEVKIDNSQVESGSL 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 143 CGYSTKKVHVFYFNKGNVINKDKDDETHLYTLVLRPDNTYEVKIDNSQVESGSL 202

Search completed: October 21, 2001, 03:07:44
Job time: 339 sec

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:08:24 ; Search time 28.98 seconds
(without alignments)
72.104 Million cell updates/sec

Title: US-09-807-148-5
Perfect score: 327
Sequence: 1 CGPGTKVHVIFNYKGNVL.....PDNTYEVKIDNSQVSGSL 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	100.0	416	1	CRTC_MOUSE
2	327	100.0	416	1	CRTC_RAT
3	327	100.0	417	1	CRTC_HUMAN
4	327	100.0	418	1	CRTC_RABIT
5	322	98.5	400	1	CRT1_BOVIN
6	322	98.5	421	1	CRT2_BOVIN
7	278	85.0	406	1	CRTC_DROME
8	251	76.8	388	1	RALL1_ONCVO
9	236	72.2	393	1	CRTC_SCHMA
10	219.5	67.1	395	1	CRTC_CAEEL
11	208	63.6	420	1	CRTC_CHLRE
12	193	59.0	416	1	CRTC_BETVU
13	188	57.5	421	1	CRTC_PRUAR
14	186	56.9	415	1	CRTC_RICCO
15	186	56.9	424	1	CRTC_ARATH
16	181	55.4	424	1	CRTC_DICDI
17	181	55.4	425	1	CRT1_ARATH
18	180	55.0	424	1	CRTC_ORYSA
19	178	54.4	416	1	CRTC_NICPL
20	175	53.5	424	1	CRTC3_ARATH
21	170	52.0	420	1	CRTC_MAIZE
22	169.5	51.8	401	1	CRTC_EUGGR
23	163	49.8	416	1	CRTC_BERST
24	117.5	35.9	592	1	CALX_HUMAN
25	116.5	35.6	591	1	CALX_MOUSE
26	116.5	35.6	591	1	CALX_RAT
27	115.5	35.3	593	1	CALX_CANFA
28	107.5	32.9	610	1	CALG_HUMAN
29	106.5	32.6	611	1	CALG_MOUSE
30	101.5	31.0	619	1	CALX_CAEEL
31	99	30.3	560	1	CALX_SCHPO
32	90.5	27.7	546	1	CALX_SOYBN
33	85.5	26.1	530	1	CAL1_ARATH

34	80.5	24.6	540	1	CALX_HELTU
35	80.5	24.6	551	1	CALX_PEA
36	77.5	23.7	528	1	CAX2_ARATH
37	73.5	22.5	540	1	MTAL_ACICA
38	61.5	18.8	582	1	SYN2_HUMAN
39	61	18.7	474	1	LAM3_MOUSE
40	61	18.7	592	1	LAM2_MOUSE
41	60.5	18.5	586	1	SYN2_RAT
42	59.5	18.2	360	1	PSD9_YEAST
43	58.5	17.9	322	1	FCY5_YEAST
44	57.5	17.6	126	1	RR13_PORPU
45	57	17.4	339	1	REP_BACAM

ALIGNMENTS

RESULT 1			
CRTC_MOUSE	STANDARD;	PRT;	416 AA.
ID	PI4211;		
AC	01-JAN-1990 (Rel. 13, Created)		
DT	01-JAN-1990 (Rel. 13, Last sequence update)		
DT	15-DEC-1998 (Rel. 37, Last annotation update)		
DE	CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60).		
GN	CALR.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 18-48 AND 129-161.		
RC	STRAIN=BAIB/C; TISSUE=Liver;		
RC	MEDLINE=90059955; PubMed=2583110;		
RA	Smith M.J., Koch G.L.E.;		
RT	"Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein.";		
RT	HACBP, a major calcium binding ER/SR protein.";		
RL	EMBO J. 8:3581-3586(1989).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	MEDLINE=93013037; PubMed=1398135;		
RA	Mazzarella R.A., Gold P., Cunningham M., Green M.;		
RT	"Determination of the sequence of an expressible cDNA clone encoding ERp50/calregulin by the use of a novel nested set method.";		
RL	Gene 120:217-225(1992).		
RN	[3]		
RC	SEQUENCE OF 18-38.		
RC	TISSUE=Fibroblast;		
RC	MEDLINE=95009907; PubMed=7523108;		
RA	Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;		
RT	"Separation and sequencing of familial and novel murine proteins using preparative two-dimensional gel electrophoresis.";		
RL	Electrophoresis 15:735-745(1994).		
CC	-1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND LOW AFFINITY CALCIUM-BINDING SITES.		
CC	-1- SUBUNIT: MONOMER (BY SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.		
CC	-1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X14926; CAA33053.1; -		
DR	EMBL; M92988; AAA37569.1; -		
DR	PIR; S06763; S06763.		
DR	PIR; JC1444; JC1444.		
DR	SWISS-2DPAGE; P14211; MOUSE.		
DR	MGD; MGI:88252; Calr.		

```

DR InterPro: IPR000886;
DR InterPro: IPR001580;
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN;
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 416
FT DOMAIN 18 137
FT DOMAIN 198 308
FT DOMAIN 309 416
FT DOMAIN 191 255
FT REPEAT 191 202
FT REPEAT 210 221
FT REPEAT 227 238
FT REPEAT 244 255
FT REPEAT 259 297
FT REPEAT 259 269
FT REPEAT 273 283
FT REPEAT 287 297
FT DOMAIN 351 407
FT DISULFID 137 163
FT SITE 413 416
SQ SEQUENCE 416 AA; 47994 MW; 24C03B00913408D8 CRC64;

Query Match 100.0%; Score 327; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 3.5e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNKGKNNLINKIRCKDDEFTHLTVLRPNTYEVKIDNSQVSGSL 60
DB 137 CGPGTKKHVIFNKGKNNLINKIRCKDDEFTHLTVLRPNTYEVKIDNSQVSGSL 196
QY 61 E 61
DB 197 E 197

RESULT 2
CRTC_RAT
ID CRTC_RAT STANDARD; PRT; 416 AA.
AC P18418; P10452;
DT 01-NOV-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (CALBP)
DE (CALCIUM-BINDING PROTEIN 3) (CABP3).
GN CALR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain cortex;
RX MEDLINE=90370496; PubMed=2395661;
RA Murthy K.K., Banville D., Srikanth C.B., Carrier F., Bell A.,
Holmes C., Patel Y.C.;
RT "Structural homology between the rat calreticulin gene product and
RT the Onchocerca volvulus antigen Ral-1."
RL Nucleic Acids Res. 18:4933-4933(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE=93202172; PubMed=8453984;
RA Nakamura M., Moriya M., Baba T., Michikawa Y., Yamanobe T., Arai K.,
Okinaga S., Kobayashi T.;
RT "An endoplasmic reticulum protein, calreticulin, is transported into
RT the acrosome of rat sperm."

```

```

RL Exp. Cell Res. 205:101-110(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE=95181573; PubMed=7876339;
RA Soennichsen B., Fuelekruug J., van Nguyen P., Diekmann W.,
Robinson D.G., Mieskes G.;
RT "Retention and retrieval: both mechanisms cooperate to maintain
RT calreticulin in the endoplasmic reticulum."
RL J. Cell Sci. 107:2705-2717(1994).
RN [4]
RP SEQUENCE OF 270-358 FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA Lone Y.C., Bailly A., Latruffe N.;
RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 18-29.
RX MEDLINE=91054414; PubMed=2241926;
RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,
MacLennan D.H., Meldolesi J., Pozzan T.;
RT "Calreticulin is a candidate for a calsequestrin-like function in
RT Ca2(+)-storage compartments (calciosomes) of liver and brain."
RL Biochem. J. 271:473-480(1990).
RN [6]
RP SEQUENCE OF 18-32.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Testis;
RX MEDLINE=92360010; PubMed=1497655;
RA Nakamura M., Michikawa Y., Baba T., Okinaga S., Arai K.;
RT "Calreticulin is present in the acrosome of spermatids of rat
RT testis."
RL Biochem. Biophys. Res. Commun. 186:668-673(1992).
RN [7]
RP SEQUENCE OF 18-32.
RC STRAIN-LEC; TISSUE=Liver;
RX MEDLINE=94072621; PubMed=8251535;
RA Yokoi T., Nagayama S., Kajiura R., Kawaguchi Y., Horiuchi R.,
Kamataki T.;
RT "Identification of protein disulfide isomerase and calreticulin as
RT autoimmune antigens in LEC strain of rats."
RL Biochim. Biophys. Acta 1158:339-344(1993).
CC !- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC !- LOW AFFINITY CALCIUM-BINDING SITES.
CC !- SUBUNIT: MONOMER (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC !- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC !- CAUTION: WAS ORIGINALLY (REF. 2) THOUGHT TO BE D-BETA-
CC HYDROXYBUTYRATE DEHYDROGENASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D78308; BAAL1345.1;
CC EMBL; X53363; CAA37446.1;
CC EMBL; X13702; CAA31987.1; ALT_SEQ.
CC EMBL; X79327; CAA55890.1;
CC PIR; S04867; S04867.
CC PIR; S11205; S11205.
CC PIR; S13045; S13045.
CC PIR; A49176; A49176.
CC PIR; S45036; S45036.
CC PIR; JH0819; JH0819.
CC InterPro; IPR000886;
CC InterPro; IPR001580;
CC Pfam; PF00262; calreticulin; 1.
CC PRINTS; PR00626; CALRETICULIN.
CC PROSITE; PS00014; ER_TARGET; 1.
CC PROSITE; PS00803; CALRETICULIN_1; 1.
CC PROSITE; PS00804; CALRETICULIN_2; 1.

```


DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 17
 FT CHAIN 18 416 CALRETICULIN.
 FT DOMAIN 18 197 N-DOMAIN.
 FT DOMAIN 198 308 P-DOMAIN.
 FT DOMAIN 309 416 C-DOMAIN.
 FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
 FT REPEAT 191 202 1-1.
 FT REPEAT 210 221 1-2.
 FT REPEAT 227 238 1-3.
 FT REPEAT 244 255 1-4.
 FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
 FT REPEAT 259 269 2-1.
 FT REPEAT 273 283 2-2.
 FT REPEAT 287 297 2-3.
 FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
 FT DISULFID 137 163 BY SIMILARITY.
 FT SITE 413 416 PREVENT SECRETION FROM ER.
 SQ SEQUENCE 416 AA; 47995 MW; 2E6713CED31A2970 CRC64;

Query Match 100.0%; Score 327; DB 1; Length 416;
 Best Local Similarity 100.0%; Pred. No. 3.5e-32;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVLINKIRKDDFFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
 DB 137 CGPGTKKHVIFNYKGNVLINKIRKDDFFTHLYTLIVRPDNTYEVKIDNSQVESGSL 196

QY 61 E 61
 DB 197 E 197

RESULT 3
 CRTCL_HUMAN STANDARD; PRT; 417 AA.
 AC P27797;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (52 KDA
 DE RIBONUCLEOPROTEIN AUTOANTIGEN RO/SS-A).
 GN CALR OR CRTCL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92013129; PubMed=1919005;
 RA Rokeach L.A., Haselby J.A., Melloy J.F., Smeenk R.J., Unnasch T.R.,
 RA Greene B.M., Hoch S.O.;
 RT "Characterization of the autoantigen calreticulin.";
 RL J. Immunol. 147:3031-3039(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90237213; PubMed=2332496;
 RA McCaulliffe D.P., Lux F.A., Lieu T.S., Sanz I., Hanke J., Newkirk M.M.,
 RA Bachinski L.L., Itoh Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,
 RA Capra J.D.;
 RT "Molecular cloning, expression, and chromosome 19 localization of a
 RT human Ro/SS-A autoantigen.";
 RL J. Clin. Invest. 85:1379-1391(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92139342; PubMed=1733953;
 RA McCaulliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;
 RT "The 5'-flanking region of the human calreticulin gene shares
 RT homology with the human GRP78, GRP94, and protein disulfide isomerase
 RT promoters.";
 RL J. Biol. Chem. 267:2557-2562(1992).

RN [4]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J., McCreedy P., Stilwagen S., Ramirez M., Carrano A.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 18-36.
 RX MEDLINE=92002034; PubMed=1911778;
 RA Rojiani M.V., Finlay B.B., Gray V., Dedhar S.;
 RT "In vitro interaction of a polypeptide homologous to human Ro/SS-A
 RT antigen (calreticulin) with a highly conserved amino acid sequence in
 RL the cytoplasmic domain of integrin alpha subunits.";
 RL Biochemistry 30:9859-9866(1991).
 RN [6]
 RP SEQUENCE OF 18-32.
 RX MEDLINE=90380058; PubMed=2400400;
 RA Krause K.H., Simmerman H.K.B., Jones L.R., Campbell K.P.;
 RT "Sequence similarity of calreticulin with a Ca2(+)-binding protein
 RT that co-purifies with an Ins(1,4,5)P3-sensitive Ca2+ store in HL-60
 RL cells.";
 RL Biochem. J. 270:545-548(1990).
 RN [7]
 RP SEQUENCE OF 18-28.
 RX TISSUE=Liver;
 RL MEDLINE=93162045; PubMed=1286669;
 RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
 RA Appel R.D., Hughes G.J.;
 RT "Human liver protein map: a reference database established by
 RT microsequencing and gel comparison.";
 RL Electrophoresis 13:992-1001(1992).
 RN [8]
 RP PARTIAL SEQUENCE OF 25-34; 56-62; 208-221 AND 273-278.
 RX TISSUE=Keratinocytes;
 RL MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [9]
 RP SEQUENCE OF 18-26.
 RX TISSUE=Colon carcinoma;
 RL MEDLINE=97295306; PubMed=9150948;
 RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
 RT "A two-dimensional gel database of human colon carcinoma proteins.";
 RL Electrophoresis 18:605-613(1997).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE THE RO AUTOANTIGEN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M84739; AAA51916.1; -;
 DR EMBL; M32294; AAA36582.1; -;
 DR EMBL; AD000092; AAB51176.1; -;
 DR PIR; A37047; A37047.
 DR PIR; S11475; S11475.
 DR PIR; A42330; A42330.
 DR PIR; A46452; A46452.
 DR SWISS-2DPAGE; P27797; HUMAN.
 DR Aarhus/Ghent-2DPAGE; 9401; IEF.
 DR HSC-2DPAGE; P27797; HUMAN.
 DR MIM; 109091; -;
 DR InterPro; IPR000886; -;

```

DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN.1; 1.
DR PROSITE; PS00804; CALRETICULIN.2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 417
FT DOMAIN 18 197 CALRETICULIN.
FT DOMAIN 198 308 N-DOMAIN.
FT DOMAIN 309 417 P-DOMAIN.
FT DOMAIN 309 417 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 408 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 414 417 PREVENT SECRETION FROM ER.
FT CONFLICT 35 35 MISSING (IN REF. 3).
SQ SEQUENCE 417 AA: 48141 MW: 84141 MW: BC37C3C0F1054FB2 CRC64;

Query Match 100.0%; Score 327; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 3.5e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKVVIFNYKGNVINKIRCKDDETHLYTLVRPNTVEYKIDNSQVSGSL 60
DB 137 CGPGTKKVVIFNYKGNVINKIRCKDDETHLYTLVRPNTVEYKIDNSQVSGSL 196

QY 61 E 61
DB 197 E 197

RESULT 4
CRTC_RABIT STANDARD; PRT; 418 AA.
AC P15253;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (BRP60).
GN CALR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=slow-twitch skeletal muscle;
RA MEDLINE=90094320; PubMed=2600080;
RX Fliegel L., Burns K., MacLennan D.H., Reithmeier R.A.F., Michalak M.;
RT "Molecular cloning of the high affinity calcium-binding protein
(calreticulin) of skeletal muscle sarcoplasmic reticulum.";
RL J. Biol. Chem. 264:21522-21528(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fast-twitch skeletal muscle;
RX MEDLINE=91282795; PubMed=2059224;
RA Fliegel L., Michalak M.;
RT "Fast-twitch and slow-twitch skeletal muscles express the same
isoform of calreticulin.";
RL Biochem. Biophys. Res. Commun. 177:979-984(1991).
RN [3]
RP SEQUENCE OF 18-36.

RX MEDLINE=91054414; PubMed=2241926;
RA Treves S., de Mattei M., Lanfretti M., Villa A., Green N.M.,
RA MacLennan D.H., Meldolesi J., Pozzan T.;
RT "Calreticulin is a candidate for a calsequestrin-like function in
Ca2(+)-storage compartments (calciosomes) of liver and brain.";
RL Biochem. J. 271:473-480(1990).
RN [4]
RP SEQUENCE OF 18-46.
RX MEDLINE=91201375; PubMed=2016321;
RA Milner R.E., Baksh S., Shemanko C., Carpenter M.R., Smillie L.,
RA Vance J.E., Opas M., Michalak M.;
RT "Calreticulin, and not calsequestrin, is the major calcium binding
protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic
reticulum.";
RL J. Biol. Chem. 266:7155-7165(1991).
RN [5]
RP PARTIAL SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=92002038; PubMed=1911780;
RA Guan S., Falick A.M., Williams D.E., Cashman J.R.;
RT "Evidence for complex formation between rabbit lung flavin-containing
monooxygenase and calreticulin.";
RL Biochemistry 30:9892-9900(1991).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC -!- LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05138; AAA31188.1; -.
DR PIR; A34154; A34154.
DR PIR; C33208; C33208.
DR PIR; D33208; D33208.
DR PIR; E33208; E33208.
DR PIR; F33208; F33208.
DR PIR; S13046; S13046.
DR PIR; S13047; S13047.
DR InterPro; IPR000886; -.
DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN.1; 1.
DR PROSITE; PS00804; CALRETICULIN.2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 418
FT DOMAIN 18 197 CALRETICULIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 418 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 408 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 415 418 PREVENT SECRETION FROM ER.
FT VARIANT 35 35 E -> D.
FT CONFLICT 90 90 P -> T (IN REF. 5).

```

```
SQ SEQUENCE 418 AA; 48275 MW; B6082B689DC763A6 CRC64;

Query Match 100.0%; Score 327; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 3.5e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
Db 137 CGPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 196

QY 61 E 61
Db 197 E 197

RESULT 5
CRT1_BOVIN STANDARD; PRT; 400 AA.
ID CRT1_BOVIN
AC P52193;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE CALRETICULIN, BRAIN ISOFORM 1 (CRP55) (CALREGULIN) (HACBP).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94183174; PubMed=8135753;
RA Matsuo K., Seta K., Yamakawa Y., Okuyama T., Shinoda T., Isobe T.;
RT "Covalent structure of bovine brain calreticulin.";
RL Biochem. J. 298:435-442(1994).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
DR InterPro; IPR000886; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00004; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Glycoprotein.
FT DOMAIN 1 180 N-DOMAIN.
FT DOMAIN 181 291 P-DOMAIN.
FT DOMAIN 292 400 C-DOMAIN.
FT DOMAIN 174 238 4 X APPROXIMATE REPEATS.
FT REPEAT 174 185 1-1.
FT REPEAT 193 204 1-2.
FT REPEAT 210 221 1-3.
FT REPEAT 227 238 1-4.
FT DOMAIN 242 280 3 X APPROXIMATE REPEATS.
FT REPEAT 242 252 2-1.
FT REPEAT 256 266 2-2.
FT REPEAT 270 280 2-3.
FT DOMAIN 334 390 ASP/GLU/LYS-RICH.
FT DISULFID 120 146 N-LINKED (GLCNAC. . .).
FT CARBOHYD 162 162 PREVENT SECRETION FROM ER (POTENTIAL).
FT SITE 397 400
SQ SEQUENCE 400 AA; 46381 MW; 7D4B68DFC689EEF1 CRC64;

Query Match 98.5%; Score 322; DB 1; Length 400;
Best Local Similarity 98.4%; Pred. No. 1.3e-31;
Matches 60; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
Db 120 CGPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 179

QY 61 E 61
Db 180 E 180

RESULT 6
CRT2_BOVIN STANDARD; PRT; 421 AA.
ID CRT2_BOVIN
AC P42918;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CALRETICULIN, BRAIN ISOFORM 2 PRECURSOR (CRP55) (CALREGULIN) (HACBP).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93385184; PubMed=8373827;
RA Liu N., Fine R.E., Johnson R.J.;
RT "Comparison of cDNAs from bovine brain coding for two isoforms of
RT calreticulin.";
RL Biochim. Biophys. Acta 1202:70-76(1993).
CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; LI3462; AAC37307.1; -.
DR InterPro; IPR000886; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00004; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 421 CALRETICULIN, BRAIN ISOFORM 2.
FT DOMAIN 35 201 N-DOMAIN.
FT DOMAIN 202 312 P-DOMAIN.
FT DOMAIN 313 421 C-DOMAIN.
FT DOMAIN 195 259 4 X APPROXIMATE REPEATS.
FT REPEAT 195 206 1-1.
FT REPEAT 214 225 1-2.
FT REPEAT 231 242 1-3.
FT REPEAT 248 259 1-4.
FT DOMAIN 263 301 3 X APPROXIMATE REPEATS.
FT REPEAT 263 273 2-1.
FT REPEAT 277 287 2-2.
FT REPEAT 291 301 2-3.
FT DOMAIN 366 411 ASP/GLU/LYS-RICH.
FT DISULFID 141 167 BY SIMILARITY.
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 418 421 PREVENT SECRETION FROM ER.
SQ SEQUENCE 421 AA; 48812 MW; 0257E959F71528BC CRC64;
```

```

Query Match          98.5%; Score 322; DB 1; Length 421;
Best Local Similarity 98.4%; Pred. No. 1.4e-31;
Matches 60; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVINKDKDETHLYTLVRPNTYEVKIDNSQVSGSL 60
   |||||
Db 141 CGPGTKKHVIFNYKGNVINKDKDETHLYTLVRPNTYEVKIDNSQVSGSL 200

QY 61 E 61
Db 201 E 201

RESULT 7
CRTC_DROME
ID CRTC_DROME STANDARD; PRT; 406 AA.
AC P29413; Q9VHA3;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP).
GN CRC OR CG9429.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208374; PubMed=1296819;
RA Smith M.J.;
RT "Nucleotide sequence of a Drosophila melanogaster gene encoding a
RL calreticulin homologue";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu P., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.E., Ferraz C., Ferrier W.S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pohlard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

```

```

Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE OF 91-124 AND 182-220.
RX MEDLINE=90307981; PubMed=2365822;
RA McCaulliffe D.P., Zappi E., Lieu T.S., Michalak M., Sontheimer R.D.,
RA Capra J.D.;
RT "A human RO/SS-A autoantigen is the homologue of calreticulin and is
highly homologous with onchocercal RAL-1 antigen and an aplysia
memory molecule";
J. Clin. Invest. 86:332-335(1990).
CC !- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC !- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC !- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X64461; CAA45791.1;
DR EMBL: AE003683; AAF54416.1;
DR PIR: A37158; A37158.
DR FlyBase: FBgn0005585; Crc.
DR InterPro: IPR000886;
DR InterPro: IPR001580;
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 406 CALRETICULIN.
FT CONFLICT 107 107 G -> A (IN REF. 3).
FT CONFLICT 184 184 V -> L (IN REF. 3).
SQ SEQUENCE 406 AA; 46808 MW; 65D72C69D0BEC427 CRC64;

Query Match          85.0%; Score 278; DB 1; Length 406;
Best Local Similarity 83.6%; Pred. No. 2.8e-26;
Matches 51; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVINKDKDETHLYTLVRPNTYEVKIDNSQVSGSL 60
   |||||
Db 137 CGPGTKKHVIFNYKGNVINKDKDETHLYTLVRPNTYEVKIDNSQVSGSL 196

QY 61 E 61
Db 197 E 197

RESULT 8
RALL_ONCVO
ID RALL_ONCVO STANDARD; PRT; 388 AA.
AC P11012;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RAL-1 PROTEIN PRECURSOR (41 KDA LARVAL ANTIGEN).
GN RALI.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;

```


QY 61 E 61
Db 195 E 195

```

RESULT 10
CRTC_CABEL
ID CRTC_CABEL STANDARD; PRT; 395 AA.
AC P27798;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 01-FEB-1994 (Rel. 28, Last annotation update)
DE CALRETICULIN PRECURSOR.
GN CRT-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=92329978; PubMed=1627827;
RA Smith M.J.;
RT "A C. elegans gene encodes a protein homologous to mammalian
calreticulin."
RL DNA Seq. 2:235-240(1992).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59589; CAA42159.1;
DR PIR; S25851; S25851;
DR InterPro; IPR000886;
DR InterPro; IPR001580;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 395
FT DOMAIN ? 192
FT DOMAIN ? 192
FT DOMAIN 193 301
FT DOMAIN 302 395
FT DOMAIN 186 250
FT REPEAT 186 197
FT REPEAT 205 216
FT REPEAT 222 233
FT REPEAT 229 250
FT REPEAT 254 292
FT REPEAT 254 264
FT REPEAT 268 278
FT REPEAT 282 292
FT REPEAT 332 390
FT DOMAIN 332 390
FT DISULFID 133 158
FT SITE 392 395
FT SEQUENCE 395 AA; 45616 MW; 35CA7D2EC1D56B03 CRC64;

Query Match 67.1%; Score 219.5; DB 1; Length 395;
Best Local Similarity 67.2%; Pred. No. 3.2e-19;
Matches 41; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

```

```

QY 1 CGPGTKKHVIFNYKGNVLINKDKDDEFTLTVRPDNTYEVKIDNSQVSGSL 60
   ||| :||||| ||||| : ||| : ||| ||||| : ||||| : |||||
Db 133 CGP-TRRVHILNYKGNLKIKEITCKSDLTHTLTILNSDNTYEVKIDGESAQTGS 191

QY 61 E 61
Db 192 E 192

RESULT 11
CRTC_CHIRE
ID CRTC_CHIRE STANDARD; PRT; 420 AA.
AC Q9STD3;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137C;
RA Zuppin A., Kaydamov C.;
RT "Cloning and characterization of a cDNA encoding Chlamydomonas
reinhardtii calreticulin."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ000765; CAB54526.1;
DR InterPro; IPR000886;
DR InterPro; IPR001580;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 420
FT SITE 417 420
FT SEQUENCE 420 AA; 47327 MW; DD3BA3AFBF61C9B CRC64;

Query Match 63.6%; Score 208; DB 1; Length 420;
Best Local Similarity 61.7%; Pred. No. 8.3e-18;
Matches 37; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNVLINKDKDDEFTLTVRPDNTYEVKIDNSQVSGSL 60
   ||| :||||| ||||| : ||| : ||| ||||| : ||||| : |||||
Db 140 CGYSTFKVHILTYKGNLKIKEITCKSDLTHTLTIVLPDNTYQVLDLKEVAGSL 199

RESULT 12
CRTC_BETVU
ID CRTC_BETVU STANDARD; PRT; 416 AA.
AC O81919;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)

```

01-OCT-2000 (Rel. 40, Last annotation update)
 DE CALRETICULIN PRECURSOR.
 OS Beta vulgaris (Sugar Beet).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
 OC Caryophyllales; Chenopodiaceae; Beta.
 OX NCBI_TaxID=3555;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VV-D/ZR5; TISSUE=Leaf;
 RA Viereck R.;
 RT "Nucleotide sequence from sugar beet calreticulin.";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ002057; CA005161.1; -;
 DR Mendel; 32549; Betvu;1166;32549.
 DR InterPro; IPR000886; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 416 CALRETICULIN.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 413 416 PREVENT SECRETION FROM ER (POTENTIAL).
 FT SITE 416 416 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 416 AA; 48136 MW; 565FEC3489F77CA7 CRC64;
 Query Match 59.0%; Score 193; DB 1; Length 416;
 Best Local Similarity 56.7%; Pred. No. 5.3e-16;
 Matches 34; Conservative 9; Mismatches 17; Indels 0; Gaps 0;
 QY 1 CGPGTKKHVIFNYKGNVLINKIRCKDDEFTHLVLPDNTYEVKINDNSQVESGSL 60
 DB 143 CGVSTKKVHAIFNYNDTHNLKIDKVPCTDQLTHVYTFILRPDATYSILIDNKEKQTGSL 202
 RESULT 13
 CRTC_PUAR
 ID CRTC_PUAR STANDARD; PRT; 421 AA.
 AC Q9X98;
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE CALRETICULIN PRECURSOR.
 OS Prunus armeniaca (Apricot).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosales; Prunaceae; Prunus.
 OX NCBI_TaxID=36596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BERGERON; TISSUE=Mesocarp, and Endocarp;
 RA Mbeguile-A-Mbeguile D., Fils-Lycaon B.R.;
 RT "Molecular cloning and nucleotide sequence of a calreticulin from
 RT apricot (Prunus armeniaca cv. Bergeron).";

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF134733; AAD32207.1; -;
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 421 CALRETICULIN.
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 418 421 PREVENT SECRETION FROM ER (POTENTIAL).
 SQ SEQUENCE 421 AA; 48416 MW; 4F5F94CBA6C6690 CRC64;
 Query Match 57.5%; Score 188; DB 1; Length 421;
 Best Local Similarity 55.0%; Pred. No. 2.2e-15;
 Matches 33; Conservative 9; Mismatches 18; Indels 0; Gaps 0;
 QY 1 CGPGTKKHVIFNYKGNVLINKIRCKDDEFTHLVLPDNTYEVKINDNSQVESGSL 60
 DB 142 CGVSTKKVHAIFNYNDTHNLKIDKVPCTDQLTHVYTFILRPDATYSILIDNKEKQTGSL 201
 RESULT 14
 CRTC_RICCO
 ID CRTC_RICCO STANDARD; PRT; 415 AA.
 AC P93508;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE CALRETICULIN PRECURSOR.
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97435975; PubMed=9290642;
 RA Coughlan S.J., Hastings C., Winfrey R. Jr.;
 RT "Cloning and characterization of the calreticulin gene from Ricinus
 RT communis L.";
 RL Plant Mol. Biol. 34:897-911(1997).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
CC -----
DR EMBL; U74631; AAB71420.1; -
DR EMBL; U74630; AAB71419.1; -
DR Mendel; 10452; Ricco:1166;10452.
DR InterPro; IPR000886; -
DR InterPro; IPR001580; -
DR Pfam; PF00262; Calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 415 CALRETICULIN.
FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SITE 412 415 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 415 AA; 47522 MW; D5F452E76CC7F8C CRC64;

Query Match 56.9%; Score 186; DB 1; Length 415;
Best Local Similarity 53.3%; Pred. No. 3.7e-15;
Matches 32; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIVNYKGNVLINKDRCKDDETHLYTLVRPNTYEVKIDNSQVSGSL 60
DB 138 CGYSTKKVHAILTYNGANHLIKKVPCEETDQLTHVTVLRPDATYSILIDNVKQTGSL 197

RESULT 15
ID CRT2_ARATH STANDARD; PRT; 424 AA.
AC Q38858; O80486; O04152;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN 2 PRECURSOR.
GN CRT2 OR CRTL OR T12M4.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Oji O., Kwan A.,
RA Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B.,
RA Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,
RA Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,
RA Theologis A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-174 FROM N.A.
RX MEDLINE-97303616; PubMed-9159940;
RA Nelson D.E., Glaunsinger B., Bohnert H.J.;
RT "Abundant accumulation of the calcium-binding molecular chaperone
calreticulin in specific floral tissues of Arabidopsis thaliana.";
RL Plant Physiol. 114:29-37(1997).
RN [3]
RP SEQUENCE OF 16-424 FROM N.A.
RC STRAIN=CV. LANDSBERG ERRECTA; TISSUE=Flower;
RA Benedetti C.E., Turner J.G.;
RT "Nucleotide sequence of an Arabidopsis thaliana cDNA encoding a
protein homologous to plant and animal calreticulins.";
RL (In) Plant Gene Register PK95-047.
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC003114; AAC24083.1; ALT_SEQ.
DR EMBL; U66344; AAC49696.1; ALT_INIT.
DR EMBL; U27698; AAA80652.1; -
DR HSP; P00268; ARXN.
DR SWISS-2DPAGE; Q38858; ARATH.
DR Mendel; 6599; Arath;1166;6599.
DR Mendel; 31147; Arath;1166;31147.
DR Mendel; 16506; Arath;1166;16506.
DR InterPro; IPR000886; -
DR InterPro; IPR001580; -
DR Pfam; PF00262; Calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein;
KW Multigene family.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 424 CALRETICULIN 2.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SITE 421 424 PREVENT SECRETION FROM ER (POTENTIAL).
FT CONFLICT 16 19 LWAI -> NSAR (IN REF. 3).
FT CONFLICT 155 155 P -> E (IN REF. 1).
FT CONFLICT 236 236 P -> T (IN REF. 3).
FT CONFLICT 408 408 V -> E (IN REF. 3).
SQ SEQUENCE 424 AA; 48084 MW; 514385EBAE810DD7 CRC64;

Query Match 56.9%; Score 186; DB 1; Length 424;
Best Local Similarity 55.0%; Pred. No. 3.8e-15;
Matches 33; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIVNYKGNVLINKDRCKDDETHLYTLVRPNTYEVKIDNSQVSGSL 60
DB 140 CGYSTKKVHAILTYNGANHLIKKVPCEETDQLTHVTVLRPDATYSILIDNVKQTGSL 199

Search completed: October 21, 2001, 03:08:25
Job time: 2/5 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:10:08 ; Search time 92.24 seconds
(without alignments)
87.496 Million cell updates/sec

Title: US-09-807-148-5
Perfect score: 327
Sequence: 1 CGPGTKKHVIFNYKGNVL.....PDNTYEVKIDNSQVSGSL 61

Scoring table: BLOSUM62

Gap: 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	303	92.7	318	13 Q9PTX7	Q9ptx7 lampetra re
2	290	88.7	405	5 Q26268	Q26268 aplysia cal
3	287	87.8	343	13 Q91711	Q91711 xenopus lae
4	287	87.8	411	13 Q91710	Q91710 xenopus lae
5	279	85.3	421	5 Q9U6S0	Q9u6s0 strongyloce
6	278	85.0	406	5 Q9U916	Q9u916 drosophila
7	276	84.4	410	5 Q16893	Q16893 amblyomma a
8	271	82.9	419	13 Q98984	Q98984 rana rugosa
9	268	82.0	417	13 Q9PU01	Q9pu01 brachydanio
10	258	78.9	387	5 Q97372	Q97372 dirofilaria
11	252	77.1	375	5 Q18478	Q18478 litomosoid
12	249.5	76.3	403	5 Q76961	Q76961 necator ame
13	246	75.2	350	5 Q26514	Q26514 schistosoma
14	246	75.2	396	5 Q45034	Q45034 schistosoma
15	226	69.1	321	13 Q9U5G0	Q9u5g0 eptatretus
16	188	57.5	421	10 Q43712	Q43712 zea mays (m
17	186	56.9	321	10 Q41799	Q41799 zea mays (m
18	185	56.6	412	10 Q40040	Q40040 hordeum vul
19	185	56.6	415	10 Q40041	Q40041 hordeum vul

20	183.5	56.1	427	10 Q9FYV2	Q9fyv2 pinus taeda
21	178	54.4	240	10 Q9ST29	Q9st29 solanum mel
22	178	54.4	389	10 Q40567	Q40567 nicotiana t
23	177	54.1	422	10 Q22502	Q22502 brassica na
24	142	43.4	214	4 Q9UDG2	Q9udg2 homo sapien
25	136.5	41.7	559	5 Q9NG26	Q9ng26 tritrichomo
26	119.5	36.5	291	5 Q9Y1V1	Q9y1v1 leishmania
27	118.5	36.2	401	5 Q9U9N9	Q9u9n9 trypanosoma
28	118.5	36.2	403	5 Q9XYF8	Q9xyf8 trypanosoma
29	116.5	35.6	582	5 Q76214	Q76214 schistosoma
30	116.5	35.6	582	5 Q9TFV3	Q9tfv3 schistosoma
31	114.5	35.0	582	5 Q04702	Q04702 schistosoma
32	109.5	33.5	622	13 Q98985	Q98985 rana rugosa
33	107.5	32.9	583	5 Q9VXF6	Q9vxf6 drosophila
34	107	32.7	397	5 Q94592	Q94592 leishmania
35	102.5	31.3	543	5 Q9VAL7	Q9val7 drosophila
36	102.5	31.3	545	5 Q9VAL7	Q9val7 drosophila
37	102.5	31.3	556	5 Q917S9	Q917s9 drosophila
38	102.5	31.3	605	5 Q02393	Q02393 drosophila
39	99	30.3	582	3 Q9HFC6	Q9hfc6 yarrowia li
40	96.5	29.5	215	6 Q9TSB7	Q9tsb7 canis sp. b
41	88.5	27.1	297	5 Q24097	Q24097 drosophila
42	79.5	24.3	272	4 Q16094	Q16094 homo sapien
43	78.5	24.0	428	10 Q41798	Q41798 zea mays (m
44	75.5	23.1	532	10 Q9LY26	Q9ly26 arabidopsis
45	73.5	22.5	540	2 Q53293	Q53293 acinetobact

ALIGNMENTS

RESULT 1

ID	Q9PTX7	PRELIMINARY;	PRT:	318 AA.
AC	Q9PTX7			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DE	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	CALRETICULIN (FRAGMENT).			
OS	Lampetra reissneri.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;			
OC	Petromyzontiformes; Petromyzontidae; Lampetra.			
OX	NCBI_TaxID=7753;			
RN	{1}			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20063780; PubMed=10594174;			
RA	Kuraku S., Hoshiyama D., Katch K., Suga H., Miyata T.;			
RT	"Monophyly of lampreys and hagfishes supported by nuclear DNA-coded			
RT	genes.";			
RL	J. Mol. Evol. 49:729-735(1999).			
DR	EMBL; AB025328; BAA88481.1; -;			
DR	InterPro; IPR000886; -;			
DR	InterPro; IPR001580; -;			
DR	Pfam; PF00262; calreticulin; 1.			
DR	PRINTS; PR00626; CALRETICULIN.			
DR	PRODOM; PD001866; -; 1.			
DR	PROSITE; PS00804; CALRETICULIN_2; 1.			
DR	PROSITE; PS00805; CALRETICULIN REPEAT; 3.			
DR	PROSITE; PS00014; ER_TARGET; UNKNOWN_1.			
FT	NON_TER			
SQ	SEQUENCE 318 AA; 36997 MW; C88102EA1CAC1506 CRC64;			

Query Match 92.7%; Score 303; DB 13; Length 318;
Best Local Similarity 91.8%; Pred. No. 2.2e-27;
Matches 56; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY	1	CGPGTKKHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSL 60	
DB	36	CGPGTKKHVIFNYKGNHLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNKVSGSL 95	
QY	61	E 61	

```

Db 96 E 96
RESULT 2
Q26268
ID Q26268 PRELIMINARY; PRT; 405 AA.
AC Q26268;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CALRETICULIN
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidae;
OC Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93098937; PubMed=1463604;
RA Kennedy T.E., Kuhl D., Barzilai A., Sweatt J.D., Kandel E.R.;
RT "Long-term sensitization training in Aplysia leads to an increase in
calreticulin, a major presynaptic calcium-binding protein.";
RL Neuron 9:1013-1024(1992).
DR EMBL: S51239; AAB24569.1;
DR InterPro; IPR000886;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 405 AA; 46738 MW; 14CAA201840D1D69 CRC64;

Query Match 88.7%; Score 290; DB 5; Length 405;
Best Local Similarity 85.2%; Pred. No. 9.1e-26;
Matches 52; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 60
Db 133 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 192
QY 61 E 61
Db 193 E 193

RESULT 3
Q91711
ID Q91711 PRELIMINARY; PRT; 343 AA.
AC Q91711;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RA Treves S., Zorzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous
system.";
RL Biochem. J. 287:579-581(1992).
DR EMBL; X67598; CAA47867.1;
DR InterPro; IPR001580;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Signal.
FT SIGNAL. 1 1
FT NON_TER. <1 12
FT CHAIN 13 411
FT SEQUENCE 411 AA; 48344 MW; 891DA66E00EBBEFA CRC64;

Query Match 87.8%; Score 287; DB 13; Length 411;
Best Local Similarity 88.5%; Pred. No. 2.1e-25;
Matches 54; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 60
Db 132 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 191
QY 61 E 61
Db 192 E 192

RESULT 5
Q906S0

```

```

DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
FT NON_TER. 1 343
FT NON_TER. 343 343
SQ SEQUENCE 343 AA; 40105 MW; 3E7DDAFA33B91DE1 CRC64;

Query Match 87.8%; Score 287; DB 13; Length 343;
Best Local Similarity 88.5%; Pred. No. 1.7e-25;
Matches 54; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 60
Db 109 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 168
QY 61 E 61
Db 169 E 169

RESULT 4
Q91710
ID Q91710 PRELIMINARY; PRT; 411 AA.
AC Q91710;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE CALRETICULIN PRECURSOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RA Treves S., Zorzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous
system.";
RL Biochem. J. 287:579-581(1992).
DR EMBL; X67597; CAA47866.1;
DR InterPro; IPR000886;
DR InterPro; IPR001580;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Signal.
FT SIGNAL. 1 1
FT NON_TER. <1 12
FT CHAIN 13 411
FT SEQUENCE 411 AA; 48344 MW; 891DA66E00EBBEFA CRC64;

Query Match 87.8%; Score 287; DB 13; Length 411;
Best Local Similarity 88.5%; Pred. No. 2.1e-25;
Matches 54; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 60
Db 132 CGPGTKKHVHIFNYKGNLVKDKDDETHLYTLVLRPDNTYEVKIDNSQVSGSL 191
QY 61 E 61
Db 192 E 192

RESULT 5
Q906S0

```

```

DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 406 AA; 46809 MW; 68BA49A6B81CC427 CRC64;

Query Match 85.0%; Score 278; DB 5; Length 406;
Best Local Similarity 83.6%; Pred. No. 2.2e-24;
Matches 51; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKVVHFIYFKGKNVLINKDIRCKDDFEFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
DB 137 CGPGTKKVVHFIYFKGKNHLSIKDIRCKDDVYTHFYTLIVRPDNTYEVKIDNEKVESGSL 196
QY 61 E 61
DB 197 E 197

RESULT 7
Q16893 ID Q16893 PRELIMINARY; PRT; 410 AA.
AC Q16893;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN.
DE CRT-1.
OS Amblyomma americanum.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Amblyomma.
OX NCBI_TaxID=6943;
RN [1]
RP SEQUENCE OF 49-410 FROM N.A.
RC TISSUE=Salivary Glands;
RA Jaworski D.C., Simmen F.A., Lamoreaux W.J., Coons L.B., Muller M.T.,
RA Needham G.R.;
RT "A secreted calreticulin protein in Ixodid tick (Amblyomma americanum
RT saliva.";
RL J. Insect Physiol. 41:369-375(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Glands;
RA Jaworski D.C.;
RT Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Glands;
RA Pain-Thornton J.M., Jaworski D.C., Needham G.R.;
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP EMBL; U07708; AAC79094.1; -.
DR InterPro; IPR000886; -.
DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866; -.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 410 AA; 47485 MW; 32CCB8750A17DC54 CRC64;

Query Match 84.4%; Score 276; DB 5; Length 410;
Best Local Similarity 83.6%; Pred. No. 3.9e-24;
Matches 51; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGPGTKKVVHFIYFKGKNVLINKDIRCKDDFEFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60
DB 136 CGPGTKKVVHFIYFKGKNHLSIKDIRCKDDVYTHFYTLIVRPDNTYEVKIDNEKVESGSL 195
QY 61 E 61
DB 196 E 196

```

```

RESULT 8
Q98984 ID Q98984 PRELIMINARY; PRT; 419 AA.
AC Q98984;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CALRETICULIN.
OS Rana rugosa (Frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8410;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96234004; PubMed=8654561;
RA Yamamoto S., Nakamura M.;
RT "Calnexin: its molecular cloning and expression in the liver of the
RT frog, Rana rugosa.";
RL FEBS Lett. 387:27-32(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Yamamoto S.;
RT "Strong expression of the calreticulin gene in the liver of Rana
RT rugosa tadpoles.";
RL J. Exp. Zool. 0:0-0(1996).
DR EMBL; D78589; BAA11425.1; -.
DR InterPro; IPR000896; -.
DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; -.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
DR InterPro; IPR000886; -.
SQ SEQUENCE 419 AA; 48658 MW; 2C857036769673BF CRC64;

Query Match 82.9%; Score 271; DB 13; Length 419;
Best Local Similarity 83.6%; Pred. No. 1.5e-23;
Matches 51; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNLQINKDKDETHLYTLVRPNTVEVKIDNSQVSGSL 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 CGPGTKKHVIFNYKGNLQINKDKADYVSHLYTLVRPNTVEVKIDNSQVSGNL 197

QY 61 E 61
Db 198 E 198

RESULT 9
Q9PUC1 ID Q9PUC1 PRELIMINARY; PRT; 417 AA.
AC Q9PUC1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CALRETICULIN.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubinstein A.L., Lee D., Luo R., Henion P.D., Halpern M.E.;
RT "Genes Dependent on Zebrafish cyclops Function Identified by AFLP
RT Differential Gene Expression Screen.";
RL Genesis 0:0-0(1999).
DR EMBL; AF195882; AAF13700.1; -.
DR InterPro; IPR000886; -.

```

```

DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 417 AA; 48723 MW; 2000C5B4004699B6 CRC64;

Query Match 82.0%; Score 268; DB 13; Length 417;
Best Local Similarity 82.0%; Pred. No. 3.3e-23;
Matches 50; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNLQINKDKDETHLYTLVRPNTVEVKIDNSQVSGSL 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 CGYSYTKKHVIFNYKGNLQINKDKDELTHLYTLVRPNTVEVKIDNEKVGSL 196

QY 61 E 61
Db 197 E 197

RESULT 10
O97372 ID O97372 PRELIMINARY; PRT; 387 AA.
AC O97372;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Dirofilaria immitis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Dirofilaria.
OX NCBI_TaxID=6287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99094497; PubMed=9879888;
RA Tsuji N., Morales T.H., Ozols V.V., Carmody A.B., Chandrashekar R.;
RT "Molecular characterization of a calcium-binding protein from the
RT filarial parasite Dirofilaria immitis.";
RL Mol. Biochem. Parasitol. 97:69-79(1998).
DR EMBL; AF052978; AAD03405.1; -.
DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; -.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 387 CALRETICULIN.
SQ SEQUENCE 387 AA; 44941 MW; E7741BF6AAFA5885 CRC64;

Query Match 78.9%; Score 258; DB 5; Length 387;
Best Local Similarity 77.0%; Pred. No. 4.4e-22;
Matches 47; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGPGTKKHVIFNYKGNLQINKDKDETHLYTLVRPNTVEVKIDNSQVSGSL 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135 CGPGTKKHVIFHYKGRNMWIKKDKDDVFTHLVTLVNSDNTVEVDGKESGEL 194

QY 61 E 61
Db 195 E 195

RESULT 11
O18478 ID O18478 PRELIMINARY; PRT; 375 AA.
AC O18478;

```

DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE RAL-1 PROTEIN (FRAGMENT).
 OS Litomosoides sigmodontis.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Litomosoides.
 OX NCBI_TaxID=42156;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA MacLennan K., Hoffman W.H., Taylor D.W.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 EMBL: AJ001621; CAA04877.1; -
 DR InterPro: IPR001580;
 DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS; PS00803; CALRETICULIN.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 FT NON_TER 375
 SQ SEQUENCE 375 AA; 43842 MW; 03F7642FBFF7A5B8 CRC64;

 Query Match 77.1%; Score 252; DB 5; Length 375;
 Best Local Similarity 75.4%; Pred. No. 2.1e-21;
 Matches 46; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

 QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEVKIDNSQVESGSL 60
 Db 135 CGPCTKKVHVIFNYKGNHMKDIRCKDDETHLYTLIVNSDNPYEVQIDGKESGEL 194

 QY 61 E 61
 Db 195 E 195

 RESULT 12
 ID 076961 PRELIMINARY; PRT; 403 AA.
 AC 076961;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE CALRETICULIN PRECURSOR.
 GN CRT.
 OS Necator americanus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Ancylostomatidae; Ancylostomatidae; Bunostominae; Necator.
 OX NCBI_TaxID=51031;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Pritchard D.I., Brown A., McElroy P., Loukas A., Girdwood K.,
 RA Berry C., Fullkrug R., Beck E.;
 RT "Calreticulin is a hookworm allergen."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 EMBL: AJ006790; CAA07254.1; -
 DR InterPro: IPR000886; -
 DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS; PS00803; CALRETICULIN.
 DR PRODOM; PD001866; -; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Signal; Allergen.
 FT SIGNAL 16 POTENTIAL.
 SQ SEQUENCE 403 AA; 46833 MW; 21F38B0515487B6F CRC64;

 Query Match 76.3%; Score 249.5; DB 5; Length 403;
 Best Local Similarity 78.7%; Pred. No. 4.5e-21;
 Matches 48; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEVKIDNSQVESGSL 60
 Db 134 CGP-TRKVDHIFSFKGNHMKDIRCKDDETHLYTLILNPNTYEVQIDGKESGEL 192

 QY 61 E 61
 Db 193 E 193

 RESULT 13
 ID 026514 PRELIMINARY; PRT; 350 AA.
 AC 026514;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE CALRETICULIN (FRAGMENT).
 GN RAL-1.
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosomatidae;
 OX Schistosoma.
 OX NCBI_TaxID=6182;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CHINESE;
 RA Huggins M.C., Moloney N.A.;
 RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
 EMBL: M80524; AAA29917.1; -
 DR InterPro: IPR000886; -
 DR InterPro: IPR001580; -
 DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS; PS00803; CALRETICULIN.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 350 AA; 40385 MW; 30FAB4E8BB685D1C CRC64;

 Query Match 75.2%; Score 246; DB 5; Length 350;
 Best Local Similarity 73.8%; Pred. No. 9.8e-21;
 Matches 45; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

 QY 1 CGPCTKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEVKIDNSQVESGSL 60
 Db 90 CGMATKKHIVIFNYKGNHMKDIRCKDDETHLYTLIVRPDNPYEVLDNKEVESGL 149

 QY 61 E 61
 Db 150 E 150

 RESULT 14
 ID 045034 PRELIMINARY; PRT; 396 AA.
 AC 045034;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE CALRETICULIN.
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosomatidae;
 OX Schistosoma.
 OX NCBI_TaxID=6182;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=PHILIPPINE (MINDORO);
 RA Scott J.C.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

Search completed: October 21, 2001, 03:10:08
Job time: 343 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:06:41 ; Search time 42.3 seconds
(without alignments)
23.852 Million cell updates/sec

Title: US-09-807-148-6

Perfect score: 258

Sequence: 1 NYGKNVINKDKRCKDEF.....PDNTYEVKIDNSQVSGSL 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2.6/ptodata/2/1aa/5A_COMB.pep.*

2: /cgn2.6/ptodata/2/1aa/5B_COMB.pep.*

3: /cgn2.6/ptodata/2/1aa/6A_COMB.pep.*

4: /cgn2.6/ptodata/2/1aa/6B_COMB.pep.*

5: /cgn2.6/ptodata/2/1aa/PCUS_COMB.pep.*

6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	55.4	415	4	US-08-675-816-2
2	100	38.8	593	1	US-08-296-362-2
3	68	26.4	542	4	US-08-675-816-6
4	61.5	23.8	582	3	US-08-906-865-3
5	54	20.9	405	1	US-07-829-954-2
6	54	20.9	405	1	US-07-994-423-2
7	54	20.9	405	1	US-08-421-891-2
8	53.5	20.7	279	2	US-08-701-191A-23
9	53.5	20.7	355	1	US-08-292-549-6
10	53.5	20.7	984	2	US-08-673-789-9
11	53.5	20.7	984	2	US-08-449-645A-19
12	53.5	20.7	984	2	US-08-702-367A-19
13	53.5	20.7	984	5	PCT-US95-04681-19
14	51.5	20.0	54	2	US-08-456-647B-41
15	51.5	20.0	54	2	US-08-237-401A-41
16	50.5	19.6	599	1	US-08-752-238-3
17	50.5	19.6	599	3	US-09-085-603B-3
18	50.5	19.6	599	3	US-09-031-897-7
19	50	19.4	1030	4	US-09-091-117-2
20	49.5	19.2	238	2	US-08-483-101-13
21	49.5	19.2	620	2	US-08-419-652-7
22	49.5	19.2	816	1	US-08-190-802A-54
23	49.5	19.2	826	1	US-07-638-431-2
24	49.5	19.2	826	5	PCT-US92-00018-2
25	49.5	19.2	1001	1	US-07-797-556-6
26	49.5	19.2	1001	1	US-07-943-843-2
27	49.5	19.2	1001	1	US-08-347-003-2

28	49.5	19.2	1097	1	US-07-943-843-6	Sequence 6, Appli
29	49.5	19.2	1097	1	US-08-347-003-6	Sequence 6, Appli
30	49	19.0	368	1	US-08-176-620A-15	Sequence 15, Appl
31	49	19.0	368	2	US-08-461-985-15	Sequence 15, Appl
32	49	19.0	770	1	US-08-525-654A-1	Sequence 1, Appli
33	49	19.0	771	1	US-08-525-654A-3	Sequence 3, Appli
34	49	19.0	976	2	US-08-449-645A-18	Sequence 18, Appl
35	49	19.0	976	2	US-08-702-367A-18	Sequence 18, Appl
36	49	19.0	976	5	PCT-US95-04681-18	Sequence 18, Appl
37	49	19.0	977	2	US-08-673-789-8	Sequence 8, Appli
38	48.5	18.8	181	1	US-07-862-005A-11	Sequence 11, Appl
39	48.5	18.8	184	1	US-07-862-005A-10	Sequence 10, Appl
40	48.5	18.8	197	1	US-07-914-284A-8	Sequence 8, Appli
41	48	18.6	569	2	US-08-467-822-21	Sequence 21, Appl
42	48	18.6	569	4	US-08-432-697-21	Sequence 21, Appl
43	47.5	18.4	238	2	US-08-483-101-12	Sequence 12, Appl
44	47.5	18.4	346	2	US-08-702-153-2	Sequence 2, Appli
45	47.5	18.4	346	2	US-08-702-153-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-675-816-2
; Sequence 2, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND US
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)-682-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-816-2

Query Match 55.4%; Score 143; DB 4; Length 415;
Best Local Similarity 50.0%; Pred. No. 9.1e-13;
Matches 24; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 1 NYGKNVINKDKRCKDEFTHLYTVIRPDNTYEVKIDNSQVSGSL 48
DB 150 NYNDTNHLIRKPCPEPDQTHVTVIRPDNTYEVKIDNSQVSGSL 197

```
RESULT 2
US-08-296-362-2
; Sequence 2, Application US/08296362
; Patent No. 5691306
; GENERAL INFORMATION:
; APPLICANT: Bergeron, John J.M.
; APPLICANT: Thomas, David Y.
; APPLICANT: Wada, Ikuro
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; TITLE OF INVENTION: PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
; TITLE OF INVENTION: PROTEIN PRODUCTION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,362
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deehr, Manya S.
; REGISTRATION NUMBER: 37,120
; REFERENCE/DOCKET NUMBER: 590066.401C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-296-362-2

Query Match 38.8%; Score 100; DB 1; Length 593;
Best Local Similarity 58.1%; Pred. No. 2.5e-06;
Matches 18; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 18 DEFTHTLTLVLPDNTYEVKIDNSQVSGSL 48
|: |||||: |||||: |||||: |||||:
Db 233 DKKHTLTLNPDNSFELVQDSIVNSGNL 263

RESULT 3
US-08-675-816-6
; Sequence 6, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND USES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)-622-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-675-816-6

Query Match 26.4%; Score 68; DB 4; Length 542;
Best Local Similarity 40.0%; Pred. No. 0.097;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 18 DEFTHTLTLVLPDNTYEVKIDNSQ 42
|: |||||: |||||: |||||: |||||:
Db 181 DKLTHVITAILKPDNELRLVDGEE 205

RESULT 4
US-08-906-865-3
; Sequence 3, Application US/08906865
; Patent No. 6040168
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; APPLICANT: Porton, Barbara
; APPLICANT: Kao, Hung-Teh
; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,865
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```


MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/421,891
FILING DATE: 14-APR-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/994,423
FILING DATE: 16-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/197/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 405 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-421-891-2

Query Match 20.9%; Score 54; DB 1; Length 405;
Best Local Similarity 36.7%; Pred. No. 7.1;
Matches 11; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 20 FTHLYTLVLRPDNTYEVKIDNSQVESGSLE 49
Db 109 FOHLQLFAKSDTSLEMTGMNALFLDGSLE 138

RESULT 8
US-08-701-191A-23
Sequence 23, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
APPLICANT: and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-23

Query Match 20.7%; Score 53.5; DB 2; Length 279;
Best Local Similarity 36.4%; Pred. No. 5.2;
Matches 12; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 5 KNVLINKIRCKDDFTHTLYTLVLRPDNTYEVK 37
Db 143 RNILVNQLCCVKVDF-GLTRLLDDFDGTETQ 174

RESULT 9
US-08-292-549-6
Sequence 6, Application US/08292549
Patent No. 5464938
GENERAL INFORMATION:
APPLICANT: Smith, Craig A.
APPLICANT: Goodwin, Raymond G.
TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,549
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/963,330
FILING DATE: 10/19/92
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2602-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-292-549-6

Query Match 20.7%; Score 53.5; DB 1; Length 355;
Best Local Similarity 26.0%; Pred. No. 7.1;
Matches 13; Conservative 12; Mismatches 16; Indels 9; Gaps 2;

QY 7 VLINKDI-----RCKDDEF-THTLYTLVLRPDNTYEVKIDNSQVESGS 47
Db 14 IINSIDITPEPSNGKCKDNEYKRRHLCLSCPPGTYASRLCDSKTNT 63

Query Match 20.7%; Score 53.5; DB 2; Length 984;

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,647B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-456-647B-41

Query Match 20.0%; Score 51.5; DB 2; Length 54;
Best Local Similarity 37.5%; Pred. No. 1.2;
Matches 12; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

QY 6 NVLLINKDIRCKDDETHLYTLIVRPDNTYEVK 37
Db 1 NILVQNLCCKVSDF-GLTRLLDDFDGTYETQ 31

RESULT 15
US-08-237-401A-41
; Sequence 41, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:

```

Search completed: October 21, 2001, 03:06:41
Job time: 331 sec

THIS PAGE BLANK (USPTO)

RESULT 2
JH0819

calreticulin precursor - rat
 N;Alternate names: calcium-binding protein 3
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 30-Sep-1993 #sequence.revision 20-Aug-1994 #text.change 20-Jun-2000
 C;Accession: JH0819; A49176; S11205; PC1109; S45036; S04867; S39372; A34473; S13045
 R;Nakamura, M.; Moriya, M.; Baba, T.; Michikawa, Y.; Yamamoto, T.; Arai, K.; Okinaga, S.
 Exp. Cell Res. 205, 101-110, 1993
 A;Title: An endoplasmic reticulum protein, calreticulin, is transported into the acrosome
 A;Reference number: A49176; MUID:93202172
 A;Accession: JH0819
 A;Molecule type: mRNA
 A;Residues: 1-416 <NAK>
 A;Cross-references: GB:D78308; NID:gl089798; PIDN:BAAL1345.1; PID:gl845572
 A;Accession: A49176
 A;Status: preliminary
 A;Molecule type: mRNA; protein
 A;Residues: 1-416 <NA2>
 A;Cross-references: GB:D78308; NID:gl089798; PIDN:BAAL1345.1; PID:gl845572
 A;Experimental source: Sprague-Dawley, spermatogenic cells
 A;Note: sequence extracted from NCBI backbone (NCBIN:127639, NCBI:P:127643)
 R;Murthy, K.K.; Banville, D.; Srikanth, C.B.; Carrier, F.; Holmes, C.; Bell, A.; Patel, V.
 Nucleic Acids Res. 18, 4933, 1990
 A;Title: Structural homology between the rat calreticulin gene product and the Onchocerca
 A;Reference number: S11205; MUID:90370496
 A;Accession: S11205
 A;Molecule type: mRNA
 A;Residues: 1-416 <MUR>
 A;Cross-references: EMBL:X53363; NID:g55854; PIDN:CAA37446.1; PID:g55855
 R;Nakamura, M.; Michikawa, Y.; Baba, T.; Okinaga, S.; Arai, K.
 Biochem. Biophys. Res. Commun. 186, 668-673, 1992
 A;Title: Calreticulin is present in the acrosome of spermatids of rat testis.
 A;Reference number: PC1109; MUID:92360010
 A;Accession: PC1109
 A;Molecule type: protein
 A;Residues: 18-32 <NAK>
 A;Experimental source: testis, strain Sprague-Dawley
 R;Sonnichsen, B.; Fueillekrug, J.; van Nguyen, P.; Diekmann, W.; Robinson, D.G.; Mieskes
 submitted to the EMBL Data Library, May 1994
 A;Description: Retention and retrieval: both mechanisms cooperate to maintain calreticulin
 A;Reference number: S45036
 A;Accession: S45036
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-416 <SOE>
 A;Cross-references: EMBL:X7937; NID:g488840; PIDN:CAA55890.1; PID:g488841
 R;Lone, Y.C.; Bailly, A.; Latruffe, N.
 submitted to the EMBL Data Library, December 1988
 A;Reference number: S04867
 A;Accession: S04867
 A;Molecule type: mRNA
 A;Residues: 'R', 270-358, 'AAG' <LON>
 A;Cross-references: EMBL:X13702; NID:g56055; PIDN:CAA31987.1; PID:g9930260
 A;Note: the authors designated the protein as D-beta-hydroxybutyrate dehydrogenase
 R;Yokoi, T.; Nagayama, S.; Kajiwara, R.; Kawaguchi, Y.; Horiuchi, R.; Kamataki, T.
 Biochim. Biophys. Acta 1158, 339-344, 1993
 A;Title: Identification of protein disulfide isomerase and calreticulin as autoimmune an
 A;Reference number: S39371; MUID:94072621
 A;Accession: S39372
 A;Molecule type: protein
 A;Residues: 18-23, 'X', 25-32 <VOK>
 R;Van, P.N.; Peter, F.; Soeling, H.D.
 J. Biol. Chem. 264, 17494-17501, 1989
 A;Title: Four intracisternal calcium-binding glycoproteins from rat liver microsomes wit
 tive calcium sequestering rat liver vesicles.
 A;Reference number: A34473; MUID:90008920
 A;Accession: A34473
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 18-36 <VAN>
 R;Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld
 Biochem. J. 271, 473-480, 1990
 A;Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage
 A;Reference number: S13045; MUID:91054414

A;Accession: S13045
 A;Molecule type: protein
 A;Residues: 18-29 <TRE>
 C;Superfamily: calreticulin
 C;Keywords: calcium binding; glycoprotein
 F;1-17/Domain: signal sequence #status predicted <SIG>
 F;18-416/Product: calreticulin #status experimental <MAT>
 F;204-212/Region: nuclear location signal
 F;413-416/Region: endoplasmic reticulum retention signal
 F;344/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 258; DB 2; Length 416;
 Best Local Similarity 100.0%; Pred. No. 4.9e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 49
 |||||
 DB 149 NYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 197

RESULT 3
 A37047
 calreticulin precursor - human
 N;Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding pr
 C;Species: Homo sapiens (man)
 C;Date: 10-Sep-1999 #sequence.revision 10-Sep-1999 #text.change 18-Feb-2000
 C;Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075
 R;McCaulliffe, D.P.; Yang, Y.S.; Wilson, J.; Sontheimer, R.D.; Capra, J.D.
 J. Biol. Chem. 267, 2557-2562, 1992
 A;Title: The 5'-flanking region of the human calreticulin gene shares homology with t
 A;Reference number: A42330; MUID:92129342
 A;Accession: A42330
 A;Molecule type: DNA
 A;Residues: 1-417 <MC2>
 A;Note: sequence extracted from NCBI backbone (NCBIN:78537, NCBI:P:78536)
 R;McCaulliffe, D.P.; Lux, F.A.; Lieu, T.S.; Sanz, I.; Hanke, J.; Newkirk, M.M.; Bachin
 J. Clin. Invest. 85, 1379-1391, 1990
 A;Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/
 A;Reference number: A37047; MUID:90237213
 A;Accession: A37047
 A;Molecule type: mRNA
 A;Residues: 1-417 <MCC>
 A;Cross-references: GB:M32294; NID:g337486; PIDN:AAA36582.1; PID:g337487
 A;Note: the authors translated the codon GTA for residue 349 as Tyr
 R;Rokeach, L.A.; Haselby, J.A.; Meilof, J.F.; Smeenk, R.J.; Unnasch, T.R.; Greene, B.
 J. Immunol. 147, 3031-3039, 1991
 A;Title: Characterization of the autoantigen calreticulin.
 A;Reference number: A46452; MUID:92013129
 A;Accession: A46452
 A;Molecule type: mRNA
 A;Residues: 1-417 <ROK>
 A;Cross-references: GB:M84739; NID:gl79881; PIDN:AAA51916.1; PID:gl79882
 A;Note: sequence extracted from NCBI backbone (NCBIN:60749, NCBI:P:60750)
 R;Lieu, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.
 J. Clin. Invest. 82, 96-101, 1988
 A;Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence
 A;Reference number: A28812; MUID:88273610
 A;Accession: A28812
 A;Molecule type: protein
 A;Residues: 18-41 <LIE>
 A;Note: 18-Ala was also found
 R;Dupuis, M.; Schaefer, E.; Krause, K.H.; Tschopp, J.
 J. Exp. Med. 177, 1-7, 1993
 A;Title: The calcium-binding protein calreticulin is a major constituent of lytic gra
 A;Reference number: PH1525; MUID:93115648
 A;Accession: PH1525
 A;Molecule type: protein
 A;Residues: 18-27 <DUP>
 A;Experimental source: LAK cell
 R;Roijani, M.V.; Finlay, B.B.; Gray, V.; Dedhar, S.
 Biochemistry 30, 9859-9866, 1991
 A;Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (c

A:Reference number: A40346; MUID:92002034
A:Accession: A40346
A:Molecule type: protein
A:Residues: 18-34; 'R' <ROJ>
R:Krause, K.H.; Simmerman, H.K.B.; Jones, L.R.; Campbell, K.P.
Biochem. J. 270, 545-548, 1990
A:Title: Sequence similarity of calreticulin with a Ca(2+)-binding protein that co-purifies with the Golgi apparatus
A:Reference number: S11475; MUID:90380058
A:Accession: S11475
A:Molecule type: protein
A:Residues: 18-32 <KRA>
R:Lamerdin, J.; McCready, P.; Stillwagen, S.; Ramirez, M.; Carrano, A.
submitted to the EMBL Data Library, November 1996
A:Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region on human chromosome 17p11.2
A:Reference number: 222906
A:Accession: T45075
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-417 <LAM>
A:Cross-references: EMBL:AD000092; PIDN:AAB511176.1
A:Experimental source: cell line 3HL2-B; fibroblast
C:Comment: Autoantibodies specific for this protein are found in Sjogren's syndrome and C:Genetics:
A:Gene: GDB:CALR
A:Cross-references: GDB:125179; OMIM:109091
A:Map position: 19p13.3-19p13.2
A:Introns: 31/1; 65/1; 133/1; 164/3; 234/3; 272/3; 320/3; 351/3
A:Note: CRPC
C:Superfamily: calreticulin
C:Keywords: calcium binding; integrin binding
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-417/Product: calreticulin #status predicted <MAT>
F:414-417/Region: endoplasmic reticulum retention signal

Query Match 100.0%; Score 258; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.9e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYKGNVLINKDKDDFFTHLYTLIVRPNTYEVKIDNSQVSGSLE 49
|||||
Db 149 NYKGNVLINKDKDDFFTHLYTLIVRPNTYEVKIDNSQVSGSLE 197
|||||

RESULT 4
A34154
calreticulin precursor, skeletal muscle - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34154; S13047
R:Fliegel, L.; Burns, K.; MacLennan, D.H.; Reithmeier, R.A.F.; Michalak, M.
J. Biol. Chem. 264, 21522-21528, 1989
A:Title: Molecular cloning of the high affinity calcium-binding protein (calreticulin) cDNA from rabbit skeletal muscle
A:Reference number: A34154; MUID:90094320
A:Accession: A34154
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <FLI>
A:Cross-references: GB:J05138; NID:g164858; PIDN:AAA31188.1; PID:g164859
R:Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld
Biochem. J. 271, 473-480, 1990
A:Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage
A:Reference number: S13045; MUID:91054414
A:Accession: S13047
A:Molecule type: protein
A:Residues: 19-32 <TRP>
C:Superfamily: calreticulin
C:Keywords: skeletal muscle
F:1-17/Domain: signal sequence #status predicted <SIG>
F:415-418/Region: endoplasmic reticulum retention signal

Query Match 100.0%; Score 258; DB 1; Length 418;

Best Local Similarity 100.0%; Pred. No. 5e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYKGNVLINKDKDDFFTHLYTLIVRPNTYEVKIDNSQVSGSLE 49
|||||
Db 149 NYKGNVLINKDKDDFFTHLYTLIVRPNTYEVKIDNSQVSGSLE 197
|||||

RESULT 5
S43376
calreticulin, brain isoform 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Oct-1994 #sequence_revision 23-Mar-1995 #text_change 07-May-1999
C:Accession: S43376; S36801
R:Matsuoka, K.; Seta, K.; Yamakawa, Y.; Okuyama, T.; Shinoda, T.; Isobe, T.
Biochem. J. 298, 435-442, 1994
A:Title: Covalent structure of bovine brain calreticulin.
A:Reference number: S43376; MUID:94183174
A:Accession: S43376
A:Molecule type: protein
A:Residues: 1-400 <MAT>
A:Experimental source: brain
R:Liu, N.; Fine, R.E.; Johnson, R.J.
Biochim. Biophys. Acta 1202, 70-76, 1993
A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin
A:Reference number: S36799; MUID:93385184
A:Accession: S36801
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 45-63; 'E', 65-83 <LIU>
A:Experimental source: brain, clone 8.1
C:Superfamily: calreticulin
C:Keywords: calcium binding; glycoprotein
F:397-400/Region: endoplasmic reticulum retention signal
F:120-146/Disulfide bonds: #status experimental
F:162/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 98.1%; Score 253; DB 2; Length 400;
Best Local Similarity 98.0%; Pred. No. 1.9e-23;
Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYKGNVLINKDKDDFFTHLYTLIVRPNTYEVKIDNSQVSGSLE 49
|||||
Db 132 NYKGNVLINKDKDDFFTHLYTLIVRPNTYEVKIDNSQVSGSLE 180
|||||

RESULT 6
S36799
calreticulin precursor, brain isoform 2 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Dec-1993 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999
C:Accession: S36799; S36800
R:Liu, N.; Fine, R.E.; Johnson, R.J.
Biochim. Biophys. Acta 1202, 70-76, 1993
A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin
A:Reference number: S36799; MUID:93385184
A:Accession: S36799
A:Molecule type: mRNA
A:Residues: 1-421 <LIU>
A:Cross-references: GB:J13462; NID:g348693; PIDN:AAC37307.1; PID:g348694
A:Experimental source: brain, clone 9.4
A:Accession: S36800
A:Molecule type: protein
A:Residues: 35-45 <LIJ>
C:Superfamily: calreticulin
C:Keywords: calcium binding; glycoprotein
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-421/Product: calreticulin, brain isoform 2 #status predicted <MAT>
F:418-421/Region: endoplasmic reticulum retention signal
F:141-167/Disulfide bonds: #status predicted
F:183/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	69.4%;	Score 179;	DB 2;	Length 395
Best Local Similarity	67.3%;	Pred. No. 2.3e-14;		

Search completed: October 21, 2001, 03:07:44
Job time: 339 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:08:25 ; Search time 28.98 Seconds
(without alignments)
57.920 Million cell updates/sec

Title: US-09-807-148-6

Perfect score: 258

Sequence: 1 NYKGNVLINKIRCKDEF.....PDNTYEVKIDNSQVESGSLE 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing: first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	258	100.0	416	1	CRTC_MOUSE
2	258	100.0	416	1	CRTC_RAT
3	258	100.0	417	1	CRTC_HUMAN
4	258	100.0	418	1	CRTC_RABIT
5	253	98.1	400	1	CRT1_BOVIN
6	253	98.1	421	1	CRT2_BOVIN
7	209	81.0	406	1	CRTC_DROME
8	182	70.5	388	1	RALL1_ONCVO
9	182	70.5	393	1	CRTC_SCHMA
10	179	69.4	395	1	CRTC_CAEEL
11	164	63.6	420	1	CRTC_CHLRE
12	148.5	57.6	424	1	CRTC_DICDI
13	145	56.2	421	1	CRTC_PRUAR
14	144	55.8	416	1	CRTC_BETUV
15	143	55.4	415	1	CRTC_RICCO
16	143	55.4	424	1	CRT2_ARATH
17	142	55.0	420	1	CRTC_MAIZE
18	138	53.5	425	1	CRT1_ARATH
19	136	52.7	401	1	CRTC_EUGGR
20	135	52.3	416	1	CRTC_NICPL
21	134	51.9	424	1	CRTC_ORYSA
22	132	51.2	424	1	CRT3_ARATH
23	128.5	49.8	416	1	CRTC_BERST
24	101	39.1	591	1	CALX_MOUSE
25	101	39.1	591	1	CALX_RAT
26	101	39.1	592	1	CALX_HUMAN
27	100	38.8	593	1	CALX_CANFA
28	95	36.8	610	1	CALG_HUMAN
29	93.5	36.2	611	1	CALG_MOUSE
30	89	34.5	560	1	CALX_SCHPO
31	78	30.2	619	1	CALX_CAEEL
32	71	27.5	546	1	CALX_SOYBN
33	68	26.4	530	1	CAX1_ARATH

34	68	26.4	540	1	CALX_HELTU	Q39994 helianthus
35	65	25.2	528	1	CAX2_ARATH	Q38798 arabidopsis
36	64	24.8	540	1	MTAL_ACTICA	P25201 actinetobact
37	61.5	23.8	582	1	SYN2_HUMAN	Q92777 homo sapien
38	60.5	23.4	586	1	SYN2_RAT	Q63537 rattus norv
39	60	23.3	474	1	LAM3_MOUSE	P48680 mus musculu
40	60	23.3	492	1	LAM2_MOUSE	P21619 mus musculu
41	57	22.1	362	1	YCY5_YEAST	P25366 saccharomyc
42	56	21.7	385	1	DIAC_HUMAN	Q01459 homo sapien
43	56	21.7	406	1	CBG_SAISC	P50451 salmirl sci
44	56	21.7	551	1	CALX_PEA	O82709 pisum sativ
45	55	21.3	339	1	REP_BACAM	P13963 bacillus am

ALIGNMENTS

RESULT 1

ID	CRTC_MOUSE	STANDARD;	PRT;	416 AA.
AC	PI4211;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP50).			
GN	CALR.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 18-48 AND 129-161.			
RC	STRAIN=BAIB/4; TISSUE=Liver;			
RX	MEDLINE=90059955; PubMed=2583110;			
RA	Smith M.J., Koch G.L.E.;			
RT	"Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein.";			
RL	EMBO J. 8:3581-3586(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93013037; PubMed=1398135;			
RA	Mazzarella R.A., Gold P., Cunningham M., Green M.;			
RT	"Determination of the sequence of an expressible cDNA clone encoding ERp60/calregulin by the use of a novel nested set method.";			
RL	Gene 120:217-225(1992).			
RN	[3]			
RP	SEQUENCE OF 18-38.			
RC	TISSUE=Fibroblast;			
RX	MEDLINE=95009907; PubMed=7523108;			
RA	Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;			
RT	"Separation and sequencing of familial and novel murine proteins using preparative two-dimensional gel electrophoresis.";			
RL	Electrophoresis 15:735-745(1994).			
CC	1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND LOW AFFINITY CALCIUM-BINDING SITES.			
CC	2- SUBUNIT: MONOMER (BY SIMILARITY).			
CC	3- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.			
CC	4- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.			
CC	5- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X14926; CAA33053.1; -			
DR	EMBL; M92988; AAA37569.1; -			
DR	PIR; S06763; S06763.			
DR	PIR; JCI444; JCI444.			
DR	SWISS-2DPAGE; PI4211; MOUSE.			
DR	MGD; MGI:88252; Calr.			

```

DR InterPro: IPR000886;
DR InterPro: IPR001580;
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 416 CALRETICULIN.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 416 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT REPEAT 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 413 416 PREVENT SECRETION FROM ER.
SQ SEQUENCE 416 AA; 47994 MW; 24C03B00913408D8 CRC64;

Query Match 100.0%; Score 258; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRGNVLLNKIRCKDDETHLYTLVIRPDNTYEVKIDNSQVESGSL 49
Db 149 NYRGNVLLNKIRCKDDETHLYTLVIRPDNTYEVKIDNSQVESGSL 197

RESULT 2
CRQC_RAT STANDARD; PRT; 416 AA.
AC F18418; P10452;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP5) (CALREGULIN) (HACBP) (ERP60) (CALBP)
DE (CALCIUM-BINDING PROTEIN 3) (CABP3).
GN CALR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain cortex;
RX MEDLINE=90370496; PubMed=2395661;
RA Murthy K.K., Banville D., Srikanth C.B., Carrier F., Bell A.,
RA Holmes C., Patel Y.C.;
RT "Structural homology between the rat calreticulin gene product and
RT the Onchocerca volvulus antigen Ral-1."
RL Nucleic Acids Res. 18:4933-4933(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE=9320172; PubMed=8453984;
RA Nakamura M., Moriya M., Baba T., Michikawa Y., Yamanobe T., Arai K.,
RA Okinaga S., Kobayashi T.;
RT "An endoplasmic reticulum protein, calreticulin, is transported into
RT the acrosome of rat sperm."
RL Exp. Cell Res. 205:101-110(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Liver;

```

```

RX MEDLINE=95181573; PubMed=7876339;
RA Soennichsen B., Fueflekrug J., van Nguyen P., Diekmann W.,
RA Robinson D.G., Mieskes G.;
RT "Retention and retrieval; both mechanisms cooperate to maintain
RT calreticulin in the endoplasmic reticulum."
RL J. Cell Sci. 107:2705-2717(1994).
RN [4]
RP SEQUENCE OF 270-358 FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA Lone Y.C., Bailly A., Latruffe N.;
RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 18-29.
RX MEDLINE=91054414; PubMed=2241936;
RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,
RA MacLennan D.H., Meldolesi J., Pozzan T.;
RT "Calreticulin is a candidate for a calsequestrin-like function in
RT Ca2(+)-storage compartments (calciosomes) of liver and brain."
RL Biochem. J. 271:473-480(1990).
RN [6]
RP SEQUENCE OF 18-32.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Testis;
RX MEDLINE=92360010; PubMed=1497655;
RA Nakamura M., Michikawa Y., Baba T., Okinaga S., Arai K.;
RT "Calreticulin is present in the acrosome of spermatids of rat
RT testis."
RL Biochem. Biophys. Res. Commun. 186:668-673(1992).
RN [7]
RP SEQUENCE OF 18-32.
RC STRAIN-LEC; TISSUE=Liver;
RX MEDLINE=94072621; PubMed=8251535;
RA Yokoi T., Nagayama S., Kajiwara R., Kawaguchi Y., Horiuchi R.,
RA Kamataki T.;
RT "Identification of protein disulfide isomerase and calreticulin as
RT autoimmune antigens in LEC strain of rats."
RL Biochim. Biophys. Acta 1158:339-344(1993).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -!- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE D-BETA-
CC HYDROXYBUTYRATE DEHYDROGENASE.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D78308; BAA11345.1;
CC EMBL: X53363; CAA37446.1;
CC EMBL: X13702; CAA31987.1; ALT_SEQ.
CC EMBL: X79327; CAA55890.1;
CC PIR: S04867; S04867.
CC PIR: S11205; S11205.
CC PIR: S13045; S13045.
CC PIR: A49176; A49176.
CC PIR: S45036; S45036.
CC PIR: JH0819; JH0819.
CC InterPro: IPR000886;
CC InterPro: IPR001580;
CC Pfam: PF00262; calreticulin; 1.
CC PRINTS: PR00626; CALRETICULIN.
CC PROSITE: PS00014; ER_TARGET; 1.
CC PROSITE: PS00803; CALRETICULIN_1; 1.
CC PROSITE: PS00804; CALRETICULIN_2; 1.
CC PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 416 CALRETICULIN.

```



```

FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 416 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 413 416 PREVENT SECRETION FROM ER.
FT SEQUENCE 416 AA; 47995 MW; 2E6713CED31A2970 CRC64;

Query Match 100.0%; Score 258; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGKNVLLINKRCKDDETHLYTLVLRPDNTYEVKIDNSQVSGSLE 49
|||||
DB 149 NYGKNVLLINKRCKDDETHLYTLVLRPDNTYEVKIDNSQVSGSLE 197

RESULT 3
CRTC_HUMAN
ID CRTC_HUMAN STANDARD; PRT; 417 AA.
AC F27797;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (52 KDA
DE RIBONUCLEOPROTEIN AUTOANTIGEN RO/SS-A).
DE CALR OR CRTC.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92013129; PubMed=1919005;
RA Rokeach L.A., Haselby J.A., Meilof J.F., Smeenk R.J., Unnasch T.R.,
RA Greene B.M., Hoch S.O.;
RT "Characterization of the autoantigen calreticulin.";
RL J. Immunol. 147:3031-3039(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90237213; PubMed=2332496;
RA McCauliffe D.P., Lux F.A., Lieu T.S., Sanz I., Hanke J., Newkirk M.M.,
RA Bachinski L.L., Itoh Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,
RA Capra J.D.;
RT "Molecular cloning, expression, and chromosome 19 localization of a
RT human Ro/SS-A autoantigen.";
RL J. Clin. Invest. 85:1379-1391(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92129342; PubMed=1733953;
RA McCauliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;
RT "The 5'-flanking region of the human calreticulin gene shares
RT homology with the human GRP78, GRP94, and protein disulfide isomerase
RT promoters.";
RL J. Biol. Chem. 267:2557-2562(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA Lamerdin J., McCready P., Stllwagen S., Ramirez M., Carrano A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 18-36.
RX MEDLINE=92002034; PubMed=1911778;
RA Rojiani M.V., Finlay B.B., Gray V., Dedhar S.;

```



```

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93385184; PubMed=8373827;
RA Liu N., Fine R.E., Johnson R.J.;
RT "Comparison of cDNAs from bovine brain coding for two isoforms of
   calreticulin."
RL Biochim. Biophys. Acta 1202:70-76(1993).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensese@isb-sib.ch).
CC -----
EMBL; LI3462; AAC37307.1; -
DR InterPro: IPR000886; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 34 POTENTIAL..
FT CHAIN 35 421 CALRETICULIN, BRAIN ISOFORM 2.
FT DOMAIN 35 201 N-DOMAIN.
FT DOMAIN 202 312 P-DOMAIN.
FT DOMAIN 313 421 C-DOMAIN.
FT DOMAIN 195 259 4 X APPROXIMATE REPEATS.
FT REPEAT 195 206 1-1.
FT REPEAT 214 225 1-2.
FT REPEAT 231 242 1-3.
FT REPEAT 248 259 1-4.
FT DOMAIN 263 301 3 X APPROXIMATE REPEATS.
FT REPEAT 263 273 2-1.
FT REPEAT 277 287 2-2.
FT REPEAT 291 301 2-3.
FT DOMAIN 366 411 ASP/GLU/LYS-RICH.
FT DISULFID 141 167 BY SIMILARITY.
FT CARBOHYD 183 183 N-LINKED (GLCNAC... ) (POTENTIAL)..
FT SITE 418 421 PREVENT SECRETION FROM ER.
SQ SEQUENCE 421 AA; 48812 MW; 0257E959F71528BC CRC64;

Query Match          98.1%; Score 253; DB 1; Length 421;
Best Local Similarity 98.0%; Pred.No.7e-25;
Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGKNVLYNKDKRCKDEFFHTLTLLVRPDNTYEYKIDNSQVESGSLE 49
    |||||
DB 153 NYGKNVLYNKDKRCKDEFFHTLTLLVRPNNTYEYKIDNSQVESGSLE 201
    |||||

RESULT 7
CRTC_DROME
ID CRTC_DROME STANDARD; PRT; 406 AA.
AC P29413; Q9VHA3;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP).
GN CRC OR CG9429.
```

OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93208374; PubMed=1296819;
 RA Smith M.J.;
 RT "Nucleotide sequence of a *Drosophila melanogaster* gene encoding a
 RT calreticulin homologue";
 RL DNA Seq. 3:247-250(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos C., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Galbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh A.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Murphy M., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtka R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 91-124 AND 182-220.
 RX MEDLINE=90307981; PubMed=2365822;
 RA McCulliffe D.P., Zappi E., Lieu T.S., Michalak M., Sontheimer R.D.,
 RA Capra J.D.;
 RT "A human Ro/SS-A autoantigen is the homologue of calreticulin and is
 RT highly homologous with onchocercal RAL-1 antigen and an aplasia
 RT 'memory molecule'";
 RL J. Clin. Invest. 86:332-335(1990).
 CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X64461; CAA45791.1; -;
 DR EMBL: AE003683; AAF54416.1; -;
 DR PIR: A37158; A37158.
 DR ElyBase: FBgn0005585; Crc.
 DR InterPro: IPR000886; -;
 DR InterPro: IPR001580; -;
 DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS: PR00626; CALRETICULIN.
 DR PROSITE: PS00014; ER_TARGET; 1.
 DR PROSITE: PS00803; CALRETICULIN_1; 1.
 DR PROSITE: PS00804; CALRETICULIN_2; 1.
 DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum: Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 406 CALRETICULIN.
 FT CONFLICT 107 107 G -> A (IN REF. 3).
 FT CONFLICT 184 184 V -> L (IN REF. 3).
 SQ SEQUENCE 406 AA; 46808 MW; 65D72C69D0BEC427 CRC64;
 .
 Query Match 81.0%; Score 209; DB 1; Length 406;
 Best Local Similarity 79.6%; Pred. No. 2.7e-19;
 Matches 39; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 NYKGNVLNKIDIRCKDSEFTLYTLIVRPDNTYEVKIDNSOVESGSL 49
 DB 149 SYAGKNHLISKIRCKDDYTFYTLIVRPDNTYEVLIIDNEKVESGNLE 197
 .
 RESULT 8
 RALL ONCVO STANDARD; PRT; 388 AA.
 ID RALL ONCVO
 AC P11012;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE RAL-1 PROTEIN PRECURSOR (41 KDA LARVAL ANTIGEN).
 GN RAL1
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94341871; PubMed=7520419;
 RA Rokeach L.A., Zimmerman P.A., Unnasch T.R.;
 RT "Epitopes of the Onchocerca volvulus RAL1 antigen, a member of the
 RT calreticulin family of proteins, recognized by sera from patients
 RT with onchocerciasis";
 RL Infect. Immun. 62:3696-3704(1994).
 RN [2]
 RP SEQUENCE OF 53-388 FROM N.A.
 RX MEDLINE=88273584; PubMed=2455736;
 RA Unnasch T.R., Gallin M.Y., Soboslay P.T., Erttmann K.D., Greene B.M.;
 RT "Isolation and characterization of expression cDNA clones encoding
 RT antigens of Onchocerca volvulus infective larvae";
 RL J. Clin. Invest. 82:262-269(1988).
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M20565; AAA59056.1; -;
 DR PIR: A32507; A32507.
 DR InterPro: IPR001580; -;
 .

```

-----
CC EMBL; M93097; AAA29854.1; -
CC EMBL; L24159; AAA19024.1; -
CC HSP; P00268; 4RXN
CC InterPro; IPR000886; -
CC InterPro; IPR001580; -
CC Pfam; PF00262; calreticulin; 1.
CC PRINTS; PR00626; CALRETICULIN.
CC PROSITE; PS00814; ER_TARGET; 1.
CC PROSITE; PS00803; CALRETICULIN_1; 1.
CC PROSITE; PS00804; CALRETICULIN_2; 1.
CC PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
CC KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
CC SIGNAL 1 16
CC FT CHAIN 17 393
CC FT DOMAIN 189 254
CC FT REPEAT 189 200
CC FT REPEAT 209 220
CC FT REPEAT 225 236
CC FT REPEAT 243 254
CC FT DOMAIN 257 295
CC FT REPEAT 257 267
CC FT REPEAT 271 281
CC FT REPEAT 285 295
CC FT CARBOHYD 27 27
CC FT DISULFID 135 161
CC FT SITE 390 393
CC FT CONFLICT 89 90
CC FT CONFLICT 188 207
CC FT CONFLICT 378 378
CC SQ SEQUENCE 393 AA; 45397 MW; 45F59857C21940D2 CRC64;
-----
Query Match 70.5%; Score 182; DB 1; Length 393;
Best Local Similarity 71.4%; Pred. No. 7.1e-16;
Matches 35; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
-----
QY 1 NYGKNVLIINKRCKDETHFTHLYLVRPDNTYEVKIDNSVSGSLE 49
      |||||  | : | : |||||  |||||  | : |||  |||||
DB 147 NYGKNHILIKETPKCKDDLXTHLYLVNPNKYEVLVDNAKVEGSLE 195
      |||||  | : | : |||||  |||||  | : |||  |||||
-----
RESULT 10
CRTC_CAEEL
ID CRTC_CAEEL STANDARD; PRT; 395 AA.
AC P27798;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE CALRETICULIN PRECURSOR.
GN CMT-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=92329978; PubMed=1627827;
RA Smith M.J.;
RT "A.C. elegans gene encodes a protein homologous to mammalian
RL calreticulin."
RL DNA Seq. 2:235-240(1992).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC , LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed (usage by and for commercial
CC purposes is allowed).
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).

CC EMBL; X59589; CA442159.1; -
CC PIR; S25851; S25851.
DR InterPro; IPR000886; -
DR InterPro; IPR001580; -
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN ? 395 CALRETICULIN.
FT DOMAIN ? 192 N-DOMAIN.
FT DOMAIN 193 301 P-DOMAIN.
FT DOMAIN 302 395 C-DOMAIN.
FT DOMAIN 186 250 4 X APPROXIMATE REPEATS.
FT REPEAT 186 197 1-1.
FT REPEAT 205 216 1-2.
FT REPEAT 222 233 1-3.
FT REPEAT 239 250 1-4.
FT DOMAIN 224 292 3 X APPROXIMATE REPEATS.
FT REPEAT 254 264 2-1.
FT REPEAT 268 278 2-2.
FT REPEAT 282 292 2-3.
FT DOMAIN 332 390 ASP/GLU/LYS-RICH.
FT DISULFID 133 158 BY SIMILARITY.
FT SITE 392 395 PREVENT SECRETION FROM ER.
SQ SEQUENCE 395 AA; 45616 MW; 35CA7D2EC1D56B03 CRC64;

Query Match 69.4%; Score 179; DB 1; Length 395;

Best Local Similarity 67.3%; Pred. No. 1.7e-15;
Matches 33; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 NYGKNVLIINKDCKDETHLYTLIVRPDNTYEVKIDNSQVSGSL 49
DB 144 NYGKNKLIKKEITCKSDDELTHLYTLILNSDNTYEVKIDGSAQTGSL 192

RESULT 11

ID CRTCL_CHLRE STANDARD; PRT; 420 AA.
AC Q9STD3;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137C;
RA Zuppin A., Kaydamov C.;
RT "Cloning and characterization of a cDNA encoding Chlamydomonas
reinhardtii calreticulin."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

CC EMBL; AJ000765; CAB54526.1; -
CC InterPro; IPR000886; -
CC InterPro; IPR001580; -
CC Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 420 CALRETICULIN.
FT SITE 417 420 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 420 AA; 47327 MW; DD3BA3AFFB61C9B CRC64;

Query Match 63.6%; Score 164; DB 1; Length 420;

Best Local Similarity 61.7%; Pred. No. 1.5e-13;
Matches 29; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 2 YKGNVLIINKDCKDETHLYTLIVRPDNTYEVKIDNSQVSGSL 48
DB 153 YKGNVLIINKDCKDETHLYTLIVRPDNTYEVKIDNSQVSGSL 199

RESULT 12

ID CRTCL_DICDI STANDARD; PRT; 424 AA.
AC Q23858;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Dictyostellum discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Mueller-Taubenberger A., Gerisch G.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).

CC EMBL; U36937; AAB87719.1; -

CC DictyDB; DD00042; -
CC InterPro; IPR000886; -
CC InterPro; IPR001580; -
CC Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 424 CALRETICULIN.
FT DOMAIN 191 257 4 X 12 AA APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 211 222 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 246 257 1-4.

```

QY      1 NYKGKNVLIINKRCKDDETHLYTLIVRPDNTYEVKIDNSOVESGSL 48
       || || || | : | : || || | : || || | : || || | : || ||
Db     154 NNNNLIKNDVPCTDDQLTHTVTFLIRPDATYSILIDNQKQTGSL 201

RESULT 14
CRTC_BTUV STANDARD; PRT; 416 AA.
ID ID CRTC_BTUV 081919; AC
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
OC Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3555;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=VV-D/ZRS; TISSUE=Leaf;
RA Viereck R.;
RC "Nucleotide sequence from sugar beet calreticulin.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; AJ002057; CAA05161.1; -.
DR Mendel; J25549; Betvu; I166; J2549.
DR InterPro; IPR000886; -.
DR InterPro; IPR001580; -.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
DR Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 416 CALRETICULIN.
FT FT CARBOHYD 57 57 N-LINKED (GLCNAC... ) (POTENTIAL).
FT FT CARBOHYD 157 157 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SITE 413 416 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 416 AA; 48136 MW; 565FEC3489F77CA7 CRC64;

Query Match 55.8%; Score 144; DB 1; Length 416;
Best Local Similarity 52.1%; Pred. No. 5.2e-11;
Matches 25; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY      1 NYKGKNVLIINKRCKDDETHLYTLIVRPDNTYEVKIDNSOVESGSL 48
       || || || | : | : || || | : || || | : || || | : || ||
Db     155 NYNDTNHLIKNDVPCTDDQLTHTVTFLIRPDATYSILIDNQKQTGSL 202

RESULT 15
CRTC_RICCO STANDARD; PRT; 415 AA.
ID ID CRTC_RICCO AC
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:10:08 ; Search time 92.24 Seconds
(without alignments)
70.283 Million cell updates/sec

Title: US-09-807-148-6
Perfect score: 258
Sequence: 1 NYKGNVLINKDIRCKDDEF.....PNTVEVKIDNSQVSGSLE 49

Scoring table:

BLOSUM62
Gapop 10.0 -, Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL16: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_unclassified: *
13: sp_vertebrate: *
14: sp_virus: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	234	90.7	318	13	Q9PTX7
2	226	87.6	343	13	Q91711
3	226	87.6	411	13	Q91710
4	221	85.7	405	5	Q26268
5	215	83.3	417	13	Q9PUC1
6	210	81.4	419	13	Q98984
7	210	81.4	421	5	Q9U680
8	209	81.0	406	5	Q9U916
9	207	80.2	410	5	Q16893
10	204	79.1	403	5	Q76961
11	196	76.0	350	5	Q26514
12	196	76.0	396	5	Q45034
13	189	73.3	387	5	Q97372
14	183	70.9	375	5	Q18478
15	179	69.4	321	13	Q9U5G0
16	148	57.4	421	10	Q43712
17	146	56.6	321	10	Q41799
18	144	55.8	412	10	Q40040
19	144	55.8	415	10	Q40041

20	140.5	54.5	427	10	Q9FTV2
21	139	53.9	214	4	Q9UDG2
22	135	52.3	240	10	Q9ST29
23	135	52.3	389	10	Q40567
24	134	51.9	422	10	O22502
25	110	42.6	559	5	Q9NG26
26	107	41.5	397	5	Q94592
27	104	40.3	291	5	Q911V1
28	99	38.4	543	5	Q9VYP4
29	99	38.4	556	5	Q917S9
30	99	38.4	622	13	Q98985
31	97	37.6	545	5	Q9VAL7
32	97	37.6	583	5	Q9VXF6
33	97	37.6	605	5	O02393
34	96	37.2	215	6	Q9TSB7
35	95	36.8	582	5	O76214
36	95	36.8	582	5	Q9TVF3
37	94	36.4	401	5	Q9U9N9
38	94	36.4	403	5	Q9XYF8
39	94	36.4	582	5	Q04702
40	84	32.6	582	3	Q9HFC6
41	83	32.2	297	5	Q24097
42	69	26.7	272	4	Q16094
43	65	25.2	428	10	Q41798
44	64	24.8	540	2	O53293
45	64	24.8	541	2	Q56752

ALIGNMENTS

RESULT 1

Q9PTX7 ID Q9PTX7 PRELIMINARY; PRT; 318 AA.
AC Q9PTX7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Lampetra reissneri.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7753;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20063780; PubMed=10594174;
RA Kuraku S., Hoshiyama D., Katoh K., Suga H., Miyata T.;
RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded genes";
RL J. Mol. Evol. 49:729-735(1999).
DR EMBL; AB025328; BAA88481.1; -;
DR InterPro; IPR000886; -;
DR InParPro; IPR001580; -;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866; -; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 318 AA; 36997 MW; C88102EALCAC1506 CRC64;

Query Match 90.7%; Score 234; DB 13; Length 318;
Best Local Similarity 89.8%; Pred. No. 5.2e-21;
Matches 44; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTVEVKIDNSQVSGSLE 49
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 48 NYKGNHLINKDIRCKDDEYTHLYTLIVRPDQTYEVKIDNKKVSGSLE 96

RESULT 2

```

Q91711 ID Q91711 PRELIMINARY; PRT; 343 AA.
AC Q91711;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=93074997; PubMed=1445218;
RA Treves S., Zorzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous
system.";
RL Biochem. J. 287:579-581(1992).
DR EMBL; X67598; CAA47867.1;
DR InterPro; IPR001580;
DR PRFam; PF00262; calreticulin; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
FT NON_TER 1 1
FT NON_TER 343 343
SQ SEQUENCE 343 AA; 40105 MW; 3E7DDAFA33B91DE1 CRC64;

Query Match 87.6%; Score 226; DB 13; Length 343;
Best Local Similarity 89.6%; Pred. No. 5.5e-20;
Matches 43; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YKGNVNLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 49
DB 122 YKKNLQINKDIRCKDDSFTHLYTLIVRPDNTYEVKIDNSKVESGSLE 169

RESULT 3
Q91710 ID Q91710 PRELIMINARY; PRT; 411 AA.
AC Q91710;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CALRETICULIN PRECURSOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=93074997; PubMed=1445218;
RA Treves S., Zorzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous
system.";
RL Biochem. J. 287:579-581(1992).
DR EMBL; X67597; CAA47866.1;
DR InterPro; IPR001580;
DR PRFam; PF00262; calreticulin; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Signal.
FT NON_TER 1 1

```

```

FT SIGNAL <1 12 POTENTIAL.
FT CHAIN 13 411 CALRETICULIN.
SQ SEQUENCE 411 AA; 48344 MW; 891DA66E00EBEFA CRC64;

Query Match 87.8%; Score 226; DB 13; Length 411;
Best Local Similarity 89.6%; Pred. No. 6.7e-20;
Matches 43; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YKGNVNLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 49
DB 145 YKKNLQINKDIRCKDDSFTHLYTLIVRPDNTYEVKIDNSKVESGSLE 192

RESULT 4
Q26268 ID Q26268 PRELIMINARY; PRT; 405 AA.
AC Q26268;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN.
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anasipidea;
OC Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93098937; PubMed=1463604;
RA Kennedy T.E., Kuhl D., Barzilal A., Sweatt J.D., Kandel E.R.;
RT "Long-term sensitization training in Aplysia leads to an increase in
calreticulin, a major presynaptic calcium-binding protein.";
RL Neuron 9:1013-1024(1992).
DR EMBL; S51239; AAB24569.1;
DR InterPro; IPR000886;
DR InterPro; IPR001580;
DR PRFam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 405 AA; 46738 MW; 14CAA201840D1D69 CRC64;

Query Match 85.7%; Score 221; DB 5; Length 405;
Best Local Similarity 81.6%; Pred. No. 2.7e-19;
Matches 40; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYKGNVNLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 49
DB 145 NYKGNVNLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNEKESGDL 193

RESULT 5
Q9PUC1 ID Q9PUC1 PRELIMINARY; PRT; 417 AA.
AC Q9PUC1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubinstein A.L., Lee D., Luo R., Henion P.D., Halpern M.E.;
RT "Genes dependent on zebrafish cyclops Function Identified by AFLP
Differential Gene Expression Screen.";

```

RESULT 7
Q9U6S0
ID Q9U6S0
PRELIMINARY;
PRT: 421 AA.

Query Match	81.0%	Score 209;	DB 5;	Length 406;
-------------	-------	------------	-------	-------------

張

Matches	35;	Conservative	5;	Mismatches	9;	Indels	0;	Gaps	0;
QY	1	NYKGNVLINKRCKDDETHLYTLVRPONTVEVKIDNSQVESGSLE	49						
		: : : : : : : : : :							
Db	147	HYKGRNHMIKKDRCDDVETHLYTLVNSDNTYEYVIDGEKAESGELE	195						
RESULT	14								
O18478		PRELIMINARY;		PRT;	375	AA.			
ID	O18478;								
AC	O18478;								
DT	O1-JAN-1998	(TEMBLrel. 05, Created)							
DT	O1-JAN-1998	(TEMBLrel. 05, Last sequence update)							
DT	O1-OCT-2000	(TEMBLrel. 15, Last annotation update)							
DE	RAL-1	PROTEIN (FRAGMENT).							
OS	Litomosoides	sigmodontis.							
OC	Eukaryota;	Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;							
OC	Onchocercidae;	Litomosoides.							
NCBI_Taxid=42156;									
OX									
RN	[1]								
SEQUENCE	FROM N.A.								
RA	MacLennan K., Hoffman W.H., Taylor D.W.;								
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AJ001621; CAA04877.1; -.								
DR	InterPro; IPR001580; -.								
DR	Pfam; PF00262; calreticulin.1.								
DR	PRINTS; PR00626; CALRETICULIN.								
DR	PROSITE; PS00803; CALRETICULIN_1; 1.								
DR	PROSITE; PS00804; CALRETICULIN_2; 1.								
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.								
NON_TER	375								
SEQUENCE	375 AA;	43842 MW;	03F7642FBFF7A5B8	CRC64;					

148 NYKGONHLIKKIDIPCKDDOKTHLYTLIVRPDNSYEVLDNKEKVESGILLE 196

RESULT	13	
O97372		
ID	PRELIMINARY;	PRT; 387 AA.
AC	O97372:	
DT	01-MAY-1999	(TREMBLrel. 10, Created)
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)
DE	01-JUN-2000	(TREMBLrel. 14, Last annotation update)
DE	CALRETICULIN PRECURSOR.	
OS	Dirofilaria immitis.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;	
OC	Onchocercidae; Dirofilaria.	
OX	NCBI_TaxID=6287;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99094497; PubMed=9879888;	
RA	Tsuji N., Morales T.H., Ozols V.V., Carmody A.B., Chandrashekar R.;	
RT	"Molecular characterization of a calcium-binding protein from the	
RT	filarial parasite Dirofilaria immitis";	
RL	Mol. Biochem. Parasitol. 97:69-79(1998).	
DR	EMBL; AF052978; AAC03405.1; ..	
DR	InterPro; IPR001580; ..	
DR	Pfam; PF00262; calreticulin; 1.	
DR	PRINTS; PR00626; CALRETICULIN.	
DR	ProDom; PD001866; -; 1.	
DR	PROSITE; PS00803; CALRETICULIN_1; 1.	
DR	PROSITE; PS00804; CALRETICULIN_2; 1.	
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.	
KW	Signal.	
FT	SIGNAL	1 18 POTENTIAL.
FT	CHAIN	19 387 CALRETICULIN.
SQ	SEQUENCE	387 AA; 44941 MW; E7741BF6AAFA5885 CRC64;

Query Match	73.3%;	Score 189;	DB 5;	Length 387;
Best Local Similarity	71.4%;	Pred. NO. 2.2e-15;		

Query Match 73.3%; Score 189; DB 5; Length 387;
Best Local Similarity 71.4%; Pred. No. 2.2e-15;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 16:49:11 ; Search time 2612.62 Seconds
 (without alignments)
 11592.138 Million cell updates/sec

Title: US-09-807-148-7
Perfect score: 1958
Sequence: 1 gaattccgcatcttctaac.....aaaaaaaaaagggaattc 1958

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:**
1: gb_ba1:**
2: gb_ba2:**
3: gb_ba3:**
4: gb_in1:**
5: gb_in2:**
6: gb_in3:**
7: gb_om:**
8: gb_ov:**
9: gb_pat1:**
10: gb_pat2:**
11: gb_ph:**
12: gb_p11:**
13: gb_p12:**
14: gb_p13:**
15: gb_p14:**
16: em_ba1:**
17: em_ba2:**
18: em_fun:**
19: em_htgo_hum:**
20: em_htgo_inv:**
21: em_htgo_rod:**
22: em_htg_hum1:**
23: em_htg_hum2:**
24: em_htg_hum3:**
25: em_htg_hum4:**
26: em_htg_hum5:**
27: em_htg_hum6:**
28: em_htg_hum7:**
29: em_htg_hum8:**
30: em_htg_inv1:**
31: em_htg_inv2:**
32: em_htg_other:**
33: em_htg_rod:**
34: em_hum1:**
35: em_hum2:**
36: em_hum3:**
37: em_hum4:**
38: em_hum5:**
39: em_hum6:**
40: em_hum7:**
41: em_in:**
42: em_om:**
43: em_or:**

44: em_ov:**
45: em_pat:**
46: em_ph:**
47: em_pl:**
48: em_ro:**
49: em_sts:**
50: em_sy:**
51: em_un:**
52: em_vi:**
53: gb_sts1:**
54: gb_sts2:**
55: gb_sts3:**
56: gb_sy:**
57: gb_un:**
58: gb_vil:**
59: gb_v12:**
60: gb_vt1:**
61: gb_vt2:**
62: gb_vt3:**
63: gb_vt4:**
64: gb_vt5:**
65: gb_vt6:**
66: gb_vt7:**
67: gb_vt8:**
68: gb_vt9:**
69: gb_vt10:**
70: gb_vt11:**
71: gb_vt12:**
72: gb_vt13:**
73: gb_vt14:**
74: gb_vt15:**
75: gb_vt16:**
76: gb_vt17:**
77: gb_vt18:**
78: gb_vt19:**
79: gb_vt20:**
80: gb_vt21:**
81: gb_vt22:**
82: gb_vt23:**
83: gb_vt24:**
84: gb_vt25:**
85: gb_vt26:**
86: gb_vt27:**
87: gb_vt28:**
88: gb_vt29:**
89: gb_vt30:**
90: gb_vt31:**
91: gb_vt32:**
92: gb_vt33:**
93: gb_vt34:**
94: gb_vt35:**
95: gb_vt36:**
96: gb_vt37:**
97: gb_vt38:**
98: em_ba3:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1958	100.0	1958	97 HUMCALRET	M84739 Human autoa
2	1837.2	93.8	1890	10 I09500	I09500 Sequence 4
3	1837.2	93.8	1890	97 HUMROSSAA	M32294 Human Ro ri
4	1831.8	93.6	1880	91 BC002500	BC002500 Homo sapi
5	1272.6	65.0	2114	94 MMCRP55	X14926 Mouse mRNA
6	1269.8	64.9	1873	94 MUSERP60A	M92988 Mouse calre
7	1268.4	64.8	1856	94 BC003453	BC003453 Mus muscu
8	1207.4	61.7	1882	95 RNCALRET	X53363 Rat mRNA fo

Db	781	ATCGATGATCCCACAGACTCCAAAGCCTGAGGACTGGGCACAAGCCGACGATATCCCTGCAC	840
Qy	841	cctgatctaagaagcccaggactggatgaagatggaacgagatggaaacccccca	900
Db	841	CCTGATCCTAAGAAGCCGAGGACTGGGATGAAGAGTGGACGGAGTGGSAAACCCECA	900
Qy	901	gtattcagaaccctgagtacaagggtgagtgaagcccccgcagatcgacaaccccagat	960
Db	901	GTCATTGAGAACCTCGTAGTACAAGGTGAGTGGAAAGCCCGCAGATCGACAACCCAGAT	960
Qy	961	tacaaggcacttggatccaccagaaaattgacaaccccgagattctccogattcccagtt	1020
Db	961	TACAAAGGCCACTTTGGATCCACCAGAAATTGACAACCCCGAGTATCTCCGATFCCAGT	1020
Qy	1021	atctatgcctatgataaactttggcttgctggccctggacctctggcaggtcaagtctggc	1080
Db	1021	ATCTATGCCCTATGATAACTTTGGCTGTGGCCCTGGACCTCTGGCAGGTCAAGTCTGGC	1080
Qy	1081	accattttgacaacttcctcatccaacacgatagggccatacgcctgaggagtttggcaac	1140
Db	1081	ACCATCTTTGACAACTTCTCATCACCAACGATGAGGCATACGCTGAGGAGTTTGGCAAC	1140
Qy	1141	gagacgtggggcgttaacaaaggcagcagagaaaaataagaaggacaacaggacagagag	1200
Db	1141	GAGACGTGGGGCTTAACAAAGCCACAGAAACAAAATGAAGGACAACAGGACGAGGAG	1200
Qy	1201	cagagccttaaggaggaggaagaagacaagaacccaagaggagagagagcagagagac	1260
Db	1201	CAGAGCCTTAAGGAGGAGGAGNAGANACANAAGAACCCAAGAGGAGGAGGACAGAGAC	1260
Qy	1261	aaggaggatgatgaggacaaagatgagtaggaggatgaggaggacaaaggaggaagat	1320
Db	1261	AAGGAGGATGATGAGGACAAAGATGAGGATGAGGAGGATGAGGAGGACAAGGAGGAAGAT	1320
Qy	1321	gaggaggaaagtccccggccaggccaaggacgagctgtagagaggcctgcctccaggg	1380
Db	1321	GAGGAGGAAGATGTCCCCGGCCAGGCCAAGGACGAGCTGTAGAGAGGCTGCCCTCCAAGG	1380
Qy	1381	ctggactgagcctgagcgctcctcgccgacagcttgcgcgcacaaataatgtctctgtg	1440
Db	1381	CTGGACTGAGCCCTGAGCGCTCCTGCCGCAGAGCTTGCCGGCCGCAATATGTCTCTGTG	1440
Qy	1441	agactcgagaacttcatTTTTTccaggctggttcggatttggggtygatttgggtttt	1500
Db	1441	AGACTCGAGAACTTTCATTTTTTTCAGGCTGGTTCGGATTTGGGTGGATTTTGGTTTT	1500
Qy	1501	gtccccctcccaactccccaccacctcccgcccttttttttatttttttaaac	1560
Db	1501	GTTCCTCTCTCCACATCTCCCCACCCCTCCCCGCCCTTTTTTTTATTTTTTAAACT	1560
Qy	1561	gglattttatctttgatctcttcagccctcacccctggcttctcatcttcttgatcaa	1620
Db	1561	GGTATTTTATCTTTGATTCTOCTTACGCCCTCACCCCTGGTTCTCACTTCTTGTATCAA	1620
Qy	1621	catettttgtcctgtcccccttctcttcttatagctccccctccaaactctggggggc	1680
Db	1621	CATCTTTTCTTGCTCTGTCCCTCTCTCATCTCTTAGCTCCCTCCAACTGGGGGGC	1680
Qy	1681	agigggttgagaagccacagccctgagatttcatctgtctctctctctctctctctct	1740
Db	1681	ASTGGTGTGGAGAAGCCACAGGCCCTGAGATTTCAFTCTCTCTCTCTCTGGAGCCAGAG	1740
Qy	1741	gaggcgacgaaggggggtgtgtcccaacccccccagcactgaggaagacgggctct	1800
Db	1741	GAGGGCAGCAGAAGGGGGTGGTGTCTCCNAACCCCCCAGCATGTAGGAAGAACGGGGCTCT	1800
Qy	1801	tctcatttccccctcccttctccccctgccccaggaactgggcaacttctgggtggggc	1860
Db	1801	TCTCATTTTACCCTCTCCCTTCTCCCTCTGCCCTCCAGGACTGGGCCACTTCTGGGTGGSGC	1860
Qy	1861	agtgggtcccagatgggtcacaactgagaatgtagaactacaacaaaaattctctataa	1920
Db	1861	AGTGGGTCCAGATTGGCTCACACTGAGAATGTAGAATGTAGAATGTACAACAAAAATTTCTATTAA	1920

Qy	1921	attaaatcttgcgtcacaataaaaaaaaagggaatttc	1958
I			
Db	1921	ATTTAAATTTTGTCTCAAAAAAAGGAATTTC	1958
I			
RESULT	2		
LOCUS	I09500	1890 bp	PAT
DEFINITION	Sequence 4 from Patent WO 8909273.		
ACCESSION	I09500		
VERSION	I09500.1	GI:587793	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1890)		
AUTHORS	Sontheimer,R.D., Lieu,T-S., Capra,J.D. and McCauliffe,D.P.		
TITLE	METHODS AND COMPOSITIONS INCORPORATING AUTOIMMUNE ANTIGENIC EPITOPES		
JOURNAL	Patent: WO 8909273-A 4 05-OCT-1989;		
FEATURES	Location/Qualifiers		
source	1..1890		
	/organism="unknown"		
BASE COUNT	464 a 497 c 520 g	409 t	
ORIGIN			
	Query Match	93.8%; Score 1837.2; DB 10: Length 1890;	
	Best Local Similarity	99.6%; Pred. No. 0;	
	Matches 1884; Conservative	0; Mismatches	3; Indels 5; Gaps
Qy	42	cgcgtactcgagacgctgcggagggtcgcttttaaaggccgcgcgttcgcccccctc	101
I			
Db	1	CCGTACTCGAGAGCGCTGCCGGAGGTCGTTTTAAAGGC--CGCGTTCCGCCGCCCTC	58
I			
Qy	102	gg-cgcgcgatcgtctaccgtgcctgtcgtcgcctcctcgcctgcgcgtgcgcgcg	160
I			
Db	59	GCCCCGCATCGCTATCCGTGCGCTGCTGCTCGGCCCTCCTGGCGCTGGCGCGT	118
I			
Qy	161	agcctgcgcgtctaactcaaggagcagttcttgagcgagacgggtggacctccccgtgga	220
I			
Db	119	AGCCCGCGCTCTACTTCAAGGAGCAGTTTCTTGACGGAGACGGGTGGACTTCCCGCTGGA	178
I			
Qy	221	tccaatccaaacacaaagtcaagattttggcaaatcgtcttcagttccgcgaagtctaacg	280
I			
Db	179	TCGAATCCAACAACAAGTCAGATTTGGCAAAATTCGTTCAGTTCGGCAAGTTCACG	238
I			
Qy	281	gtcacgaggaagaagataaaggtttgcagacaagccaggatgcacgcttttatgctctgt	340
I			
Db	239	GTGACGAGAGAAGATAAAGSTTTGCAGACAACCAGGATGCACGCTTTATGCTCTGT	298
I			
Qy	341	cggccagtttcgagccttcagcaacaaagccagacgcgttgttgtagttccagttcacggtga	400
I			
Db	299	CGGCCAGTTTCGAGCCTTTCAGCAACAAGGCCAGACGCTGTTGGTGCAGTTTCACGGTGA	358
I			
Qy	401	aacatgaacaaacatacgaactgtggggcgctatgtgaagctgtttccctaagtatttg	460
I			
Db	359	AACATGACGAACAATCGACHTGTGGGGCGGCTATGTGAAGCTGTTCCTTAATAGTTTGG	418
I			
Qy	461	accagacagacatgcacggagactcagaataacacatcatgtttgggtcccgacatctgtg	520
I			
Db	419	ACCAGACAGATGCACGGGACACTCAGAATAACACATCATGTTTGGTCCCACATCTGTG	478
I			
Qy	521	gccctggcaccagaaggttcattgcattcttaactacaaggggcaagaacgtgcgtatca	580
I			
Db	479	GCCCTGGCACCAAGAAGGTTTCATGTCATCTTCAAATCTACAAGGCAAGAAGCTGCTGATCA	538
I			
Qy	581	acaaggacatccgttgaagatgatgaatttacacacctgtacacactaatgtgcgcg	640
I			
Db	539	ACAAGGACATCCGTTGCAAGGATGATGAGTTTACACACCTGTACACACTGATGTTGCGCG	598
I			
Qy	641	caqacaaacacctatgagttgaagattgaacaacgccagggtggagtcgcggtcccttggag	700


```
Db 59 GGGCCGCATGCTATCCGTCGCGTGTGCTCGGCCCTCCTCGGCGTGGCGTCGCG 118
Qy 161 agctgcctctacttcaagagcagttcttgacgagacgggtggaacttcccgctgga 220
Db 119 AGCCCGCGTCTACTTCAAGGAGCAGTTTCTGACGGAGACGGGTGGACTTCCCGCTGGA 178
Qy 221 tcgaatccaaacaacagtcagatttggcaaatcgcttccagttcccggaagctctacg 280
Db 179 TCGAATCCAAACAACAGTCAGATTTTGGCAAAATCGTTCTCAGTTCGGCGCAAGTCTACG 238
Qy 281 gtacagagggaagataaagggtttgcagacaagccaggatgcacgcgttttatgctctgt 340
Db 239 GTGACGAGGAAAGATAAAGGTTTGACAGAAAGCCAGGATGCACGCTTTTATGCTCTGT 298
Qy 341 cgccagtttcagagcttcagacaacaaagccagacgctgggtgcagttcacgtgga 400
Db 299 CGCCAGTTTCGAGCCCTTCAGCAACAAGCCAGACGCTGGTGGTGCAGTTTCACGGTGA 358
Qy 401 aacatgagcagaacatcgactgtggggcggtctatgtgaagctgtttccctaatagtttg 460
Db 359 AACATGAGCAGACATCGACTGTGGGGCGGCTATGTGAAGCTGTTCCTAATAGTTTG 418
Qy 461 accagacagacatgcacggagatcagaaatacaacatcatgtttggtcccgacatctgtg 520
Db 419 ACCAGACAGACATGCACGGAGACTCAGAATACAACATCATGTGTGGTCCGACATCTGTG 478
Qy 521 gccctggcaccagaaggttcactatcttcaactacaaggccaagaacgctgctgaca 580
Db 479 GCCTTGGCACCAAGAGGTTTCATGTCTATCTTCAACTACAAGGGCAAGAAGCTGCTGATCA 538
Qy 581 acaaggacatccgtgtcaaggatgatgtttacacacctgttacacactgattgtcggc 640
Db 539 ACAAGGACATCCGTTGCAAGGATGATGAGTTTACACACCTGTACACACTGATTGTGGCGC 598
Qy 700 cagacaacacctatgaggtgaagtgtacaacagccaggtgaggtccggtcccttgaag 700
Db 599 CAGACAACACCTATGAGGTGAAGATTGACAACAGCCAGGTGGAGTCCGGCTCCTTGGAAAG 658
Qy 701 acgatggagactcctgcacccaagaagataaaggatcctgatgcttcaaaacccgaag 760
Db 659 ACGATTGGGACTTCTCTGCCACCCAAAGAGATAAAGGATCCTGATGCTCAAAAACCGGAAG 718
Qy 761 actggatgagcgggcacaagatcgatgccacagactccaagcctgagactgggaca 820
Db 719 ACTGGGATGAGGGGCCAAGATCGATGATCCACAGACTCCAAGCCTGAGGACTGGGACA 778
Qy 821 agcccagcataccctgacccctgatgctaagaagcccgaggactgggtgaagagatgg 880
Db 779 AGCCGAGCATATCCCTGACCCCTGATGCTAAGAAGCCCGAGGACTGGGATCAAGAGATGG 838
Qy 881 acggagagtgggaacccccagtgattcagaacccctgagtaacaagggtgagtgaagcccc 940
Db 839 ACCGAGAGTGGGAACCCCAAGTATTGAGAACCCCTGAGTACAAAGGTGAGTGAAGCCCC 898
Qy 941 ggcatatcgacaacccagattacaaggcacttggatccaccagaaattgacaaccccc 1000
Db 899 GGCAGATCGAACCCCAAGATTACNAGGGCACTTGGATGCCACCAGAAATTGACAACCCCG 958
Qy 1001 agtatctcccgatcccgatctatcgctatgataaacttggcgctgggctggacc 1060
Db 959 AGTATTCTCCCGATCCCGATGATCTATGCCATGATACTTTGGCGTGTGGCCCTGGACC 1018
Qy 1061 tctggcaggtcaagtctggcacaacttctgacaacttccctcaaccacagatgagcat 1120
Db 1019 TCTGGCAGGTCAGCTCTGGCACCATTCTTGCACAACTTCTCATCACCACAGATGAGGAT 1078
Qy 1121 acgtgagaggttggcaacgaacgtggcggtacaaaggcagcagagaaacaaatga 1180
Db 1079 ACCTGAGGAGTTTGGCAACAGACGCTGGGCGGTAAACAAAGGCGAGAGAAACAAATGA 1138
Qy 1181 aggacaacaggagcagagcagaggtcttaaggagggagggaagacaagaacgcgaag 1240
```

```
Db 1139 AGGACAAACAGGAGGAGGAGCAGAGGCTTAAGGAGGAGGAAGAAGACAAGAAACGCAAG 1198
Qy 1241 agagagagagagcagagacaagagatgatgagacaagatgagagatgagagatg 1300
Db 1199 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGATGATGAGGACAAAGATGAGGATGAGGAGGATG 1258
Qy 1301 aggaggaagaggaagatgagaggaagatgtcccccggccagggccaagacagagctgt 1360
Db 1259 AGGAGGACAAAGGAGGAGATGAGGAGGAAGATGTCCCGCGCAGGCCAAGCACGAGCTGT 1318
Qy 1361 agagagcctgcctccaggcctggaactgagcctgagcctcctgcgcgacagattgccg 1420
Db 1319 AGAGAGCCCTGCCCTCCAGGCTGGACTGAGCCCTGAGCGCTCTCCGCGACAGCTTGCCG 1378
Qy 1421 cgccaaataatgtctgtgagactcgagaactttcatttttccaggtcggttcgat 1480
Db 1379 CGCCAAATAATGCTCTGTGAGACTCGAGAACTTTCATTTTTTCCAGGCTGGTTCGAT 1438
Qy 1481 ttgggggtgattttgttttccctcctccactctcccccccccctccccccctt 1540
Db 1439 TTGGGGTGGATTTTGGTTTGTTCCTCCCTCCCTCCACTCTCCCCACCCTCCCGCCCTT 1498
Qy 1541 ttttttatttttttaaacgtgtatttat-ctttgattctccttcagcctcaccctgt 1599
Db 1499 TTTTTTTTTTTTTTTAAACTGGTATTTATCCTTTGATTTCTCTTCAGCCCTCACCCTG 1558
Qy 1600 gtctcatcttcttgatcaacatctttctgctctgt-ccccctctctcatctcta 1658
Db 1559 GTTCTCATCTTCTTGATCAACATCTTTCTTGCTCTGTGCTCTCTCTCTCTCTCTCTCTCT 1618
Qy 1659 gctccccctcaacacctggggcgagtggtgagaagccacagcctgagatttcactgt 1718
Db 1619 GCTCCCCCTCCAACTGGGGGCGAGTGTGTGGAGAAGCCACAGGCTTGAGATTTTCATCTG 1678
Qy 1719 ctctcctcctgagccagagggcgagcagagaaggggtggtgtctccaaacccccag 1778
Db 1679 CTCTCTCTCTCTGAGCCAGAGGAGGCGAGAGGGGGTGGTGTCTCTCAACCCCCAG 1738
Qy 1779 cactgagaagaacgggctctctcatcttccacccctctcttccccctgccccagga 1838
Db 1739 CACTGAGAGAAGCGGGCTCTTCTCATTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1798
Qy 1839 ctggggcactctctgggtggggcagtggtcccgatgggtccacactgagaaatgtaaga 1898
Db 1799 CTGGCCACTCTCTGGTGGGGCAGTGGGTCCAGATTTGGCTCACACTCAGAAATGTAAGAA 1858
Qy 1899 ctacaacaaattcttattataattttt 1930
Db 1859 CTACAAACAAATTTCTATTAAATTAATTTT 1890
```

RESULT 4

```
BC002500 1880 bp mRNA PRI 16-MAR-2001
LOCUS
DEFINITION Homo sapiens, calreticulin, clone MGC:2227, mRNA, complete cds.
ACCESSION BC002500
VERSION BC002500.1 GI:12803362
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1880)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
```

Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhri.nih.gov
Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Katesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.B., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, S., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Tuchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 4 Row: g Column: 3.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```
location/qualifiers
1. 1880
/organism="Homo sapiens"
/db_xref="LocusID:411"
/db_xref="taxon:9606"
/clones="MGC:2227 IMAGE:3050717"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH.MGC.20"
/lab_host="DH10B-R"
/note="vector: pORF7"
43. 1296
```

CDS

```

43. .1296
/codon_start=1
/product="calreticulin"
/protein_id="AAH02500.1"
/db_xref="GI:12803363"
/translation="MLLSVPLLLGLIG
DFGKGVSSGKFGDEERKDKGLQTSQD
IDCGGCVKLFPSNLDQTDMDHGDSEYN
IRCKDDEFTHLITLIVRPDNTYEVKID
WDERAKIDDPDTDSKPEOWDKPEHIDPP
PRQIDNPDKGTWTHPEIDNFEYSPDP
DEAYAEFGNETGWGTYKAAEQMKDKQ
EDEDDEEDKDEEDDEEDVPGQAKDEL"
482 a 4499 c 513 g 396 t
BASE COUNT

```

Query Match 93.6%; Score 1831.8; DB 91; Length 1880;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1853; Conservative 0; Mismatches 17; Indels 1;

Qy	77	aagggccgcgcttgcgcgcccccctcgg~ccgcgaatgctgtatccgtgcggtttgctctc	135
Db	10	AGGCCGCGCGTGTGGCCGCCCTTCGGCCGCATGCTGCTATCCGTGCCGCTCCTGCTC	69
Qy	136	ggcctcctcgccctggcgcgtgcgcgacgctgcgccttcacttcgaaggacagtttctctggac	195
Db	70	GSCCTCTCGGCCCTFGCCGCTCGCCGAGCCTGCCCTACTTCAAGGAGACAGTTTCTTGAC	129
Qy	196	ggagacgggttggaacctccgcctgcgatcgaatacacaacacaagaatcagattttggcaaatc	255
Db	130	GGAGACGGGTGGACTTCCCCTGGATCGAATCCAACAACAAAGTCAGATTGTGGCAAATTC	189
Qy	256	gtctcagttccgcgaagtctcacgtgcagcagagaaagataaagcttgcagacaagc	315
Db	190	GTTCTCAGTTCGGCACAGTTCTACGGTAGCAGAGAGAAAAGATAAAGGTGTTCACACAAACG	249
Qy	316	caggatgcacgcttttatgctctgcgcgcagtttcgagcctttcagcaacaagaagccag	375
Db	250	CAGGATCACGCTTTTATGCTGTGTCGCCAGTTTCAGGCCTTTCAGCAACAAGGCCAG	309
Qy	376	acgctgttgtgcagtttcacggtgaaaacatgacagacaatcatcgaactgtggggcgcgctat	435
Db	310	ACGCTGCTGTGTGAGTTTCACGGTGAACAATGAGCAGAACAATCGACTCTGGGGCGGGCTAT	369


```
QY 883 ggagagtgggaaccccaagtgattcagaacccctgagtcacaaggtgagtggaagccccc 942
Db 844 GGAGAGTGGGAACACACAGTGTATCAAAATCCCTGAATACAAAGGCGAGTGGAAACACGT 903
QY 943 cagatgacacccagattacaaggcacttgatccaccacgaataatgacaaccccgag 1002
Db 904 CAAATGACAACCCAGATTACAAGGTACCTGGATACACCACCAAAATTGACAACCCCTGAA 963
QY 1003 tattctccagatcccagatctatgctatgataaatttggcgtgctggcctggaccc 1062
Db 964 TACTCCCCGATCAAAATCTATGCTCTATGATAGTTTGTCTACTGGCCTAGATCTC 1023
QY 1063 tgcaggtcaagtctgacacactcttgacacacttctcatcaccacacatgagcatatc 1122
Db 1024 TGGCAGGTCAAGTCTGGCAAACTTTGACAAATTTCTCATCAACCAATGATGAGGCTAT 1083
QY 1123 gctgagagttgttgcaacgagacgtgggctgaacaaagcagcagagaacaaatgaag 1182
Db 1084 GCAGAGAGTTTGGCAATGAGAGCTGGGGTGTACCAAGGCTGCAGAGAAAGCATGAAG 1143
QY 1183 gacaacaggacagagagagagcgttaagagagagagagagagagagagagagagagag 1242
Db 1144 GACAAGCAGGATGAGGAGCAGAGGCTTAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1203
QY 1243 gagagagagcagagacagagagagatgatgagacaaagatgagatgagagagatgag 1302
Db 1204 GAAGAAGAGCTGAGGATTAAGAGGATGATGATGACAGAGATGAATGAGGAGAGAGAA 1263
QY 1303 gaggaacagagagaagatgagaggaagatgtcccccggcagagccacagagagctgtag 1362
Db 1264 GATGAGAAGGAGGAAGAT---GAGGAAGAATCCCTTGGCCAAAGCAAGGATGAGCTGTAG 1320
QY 1363 ag-----aggcctgctccagggctggactgagggcctgagcgtcctgccgagagcct 1416
Db 1321 AGCCACACACACCTGCTTACAGGGCTGGACTGAGGCTGAACAC-CCTGCCGAGAGAGCTG 1379
QY 1417 gccgcgcaaaatagtctctgtgagactcagaaacttcatctttttccagctggttc 1476
Db 1380 GCTGCTCCCAATTAATGCTCTATGAGACTCAAGAACTTTTTCATTTTTCAGGACAGGTTT 1439
QY 1477 ggatttgggtggtatttggtttggctccctcctcctcctcctcctcctcctcctcctc 1536
Db 1440 AGATCTGGGTAGATCTGATTTGTTGCCCTGC-----CTCCCCCATTAACCCCTCC 1491
QY 1537 ccttttttttttttttaaaactggtattttatttttttttttttttttttttttttttt 1596
Db 1492 CTTTTTTTTTTTACTGTGT-----TTTGTCTTTAATTCCTCTCAGCCCTCATCT 1540
QY 1597 ctggttctcatctttctgatacaacatcttttcttctcctctgtccctctctctcatctct 1656
Db 1541 GGTTCCTCATTTTGAATCAACATCTTTCCTCTCTCTCCCTCCCTTCTCCATCTTTTGG 1600
QY 1657 tagtccctccccaacctgggggagctggtgt---ggagaagccacagagcctgagatttc 1713
Db 1601 TCACCTACCTCCCACTCTAGGAACAGGGGTGTAGAGGAGAGAGCCCTAGGCTTGAGATTTC 1660
QY 1714 atctgtctctctcctctggagccagagagggcgagcagagaggggtggtgtctccaaacc 1773
Db 1661 ATCTGCTCTCTCTCTGCAATCTCAGAGGAGGCGAGGAGAGAGGGGTGTGTTTTCCTCTCC 1720
QY 1774 ccagcactgaggaagaagggtctctctcatcttaccctcctctcttccctcctccccc 1833
Db 1721 CCCCAGCTGAGGAAGAAATGGGGCTCTTCTCAT-----CCCCCTTCTCCCTTCCCTCC 1772
QY 1834 caggactggggcacttctgggtggggcagtggtgtccagattgggtccacactgagaatgt 1893
Db 1773 CAGGACTGGGCCACTT---GTGGGCGAGCCAGTCTTAGCACAGCTCACACTGAGAGTGT 1828
QY 1894 aagaactacaaacaaatttctattataataatttttggctct 1936
Db 1829 AAGAACTACAAACAAAATTTCTATTAAATTAAGTTTTGTGTCT 1871
```

```
RESULT 6
MUSERP60A 1873 bp mRNA ROD 27-APR-1993
LOCUS Mouse calregulin (Erp60) mRNA, complete cds.
DEFINITION M92988
ACCESSION M92988
VERSION 1
KEYWORDS GI:193084
SOURCE Mus musculus cDNA to mRNA; and Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1873)
AUTHORS Mazzarella,R.A., Gold,P., Cunningham,M. and Green,M.
TITLE Determination of the sequence of an expressible cDNA clone encoding
JOURNAL Erp60/calregulin by the use of a novel nested set method
MEDLINE Gene 120, 217-225 (1992)
FEATURES
FEATURES Location/Qualifiers
1..1873
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="fibroblast"
63..1319
/gene="Erp60"
69..1319
/gene="Erp60"
/codon_start=1
/product="calregulin"
/protein_id="AAA37569.1"
/db_xref="GI:193085"
/translaton="MLLSVPLLLGLLLAAADPAIVFKQFLDGDANTNRWVESKHS
DEGKPVLSGKEYGDLKDKGLQTSODARFYALSAKEEFSNKGOTLVVQFTVKEHQN
IDCGGYKALFFSLGDLQDKMHGSEYNMFPGDICGPGTKKVVIFNYKGNVLIINKD
ICKQDEFTHLYTLIVPONTYEKIDNSQVSGLEDWDDELPPKIKIDPKDAARKEP
WDERAKIDDPDTSKPEDPKPEHIDPDPAKPEDWDDEDEWEPPEVPIQNGEYKGEWK
PQIOPNDYKGTWIIHPEIDNPESPDANIYAYDSFAVLGLDLQVKSQGTIFNFLTIN
DEAYAEFGNETGVTKAAEKOMKQKDEQRLKEEEDKKRKEEAEDEKDDDDDD
DEDEDEDEKEDEEESPGQAKDEL"
69..1119
/sig_peptide
/gene="Erp60"
120..1316
/mat_peptide
/gene="Erp60"
1305..1316
/misc_feature
/gene="Erp60"
/note="retention signal"
BASE COUNT 497 a 463 c 485 g 428 t
ORIGIN
Query Match 64.9%; Score 1269.8; DB 94; Length 1873;
Best Local Similarity 83.1%; Pred. No. 3.4e-244;
Matches 1581; Conservative 0; Mismatches 277; Indels 45; Gaps 10;
QY 45 tactgcagagccgctccgcgaggtcgctgttttaaaaggcc-gcgcgttgcgcgccccctcg 103
Db 3 TACCGAGAGCCGCTGCCCTGAAGATCGTCTTAAAGSCCTGTGTGCCGCCGCCCTCGG 62
QY 104 -ccgcacatgctgctatccgtgcggttgcgtgcgtcctcctcctcctcctcctcctc 162
Db 63 CCGCCCATGCTCTCTTTCGGTGGCGCTCTCTTGGCCTCTCTCGGCCCTGGCGCCGAGAC 122
QY 163 cctgcgctactctcaagagacagttcttgacgagaggggtggacttcccgctgagtc 222
Db 123 CTTGCCATCTATTTTAAAGAGCAGTCTTTGGACGGAGATGCCCTGGACCAACCCCTGGGTC 182
QY 223 gaatccaaacacacacagtcagatttttgcaaatctgtctcagttcccggaagttctacgt 282
Db 183 GAATCCAAACATAAGTCCGATTTTGGCAAAATTTGCTCTCAGTTCTGGCAAAATTTACGGG 242
QY 283 gacgaggaagaagataaagtttgcagagacagccagagatgcacgcttttatgtctgtcg 342
Db 243 GACCTGGAGAAGGATAAAGGGCTGACAGACAAGCAAGATGCCCGGATTTTACGCACTGTCC 302
```


Series: IRAK Plate: 5 Row: e Column: 24.

FEATURES

source

Location/Qualifiers
1. .1856
/organism="Mus musculus"
/db_xref="LocusID:12317"
/db_xref="taxon:10090"
/clone="MGC:6209 IMAGE:2655918"
/issue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
old mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
37. .1287
CDS
/codon_start=1
/product="Galreticulin"
/protein_id="AAH03453.1"
/db_xref="GI:13097432"
/translation="MLLSVPLLLGLLGLAADPAIYFKEQFLDGDWNRWVSKHS
DFGKVLSSGKFGDLEKDLQTSQDARFALSASFPSNKGOTLVVQVTVKHEQN
IDCGGVYKLFPSGLDQDMHGDSEYINFGDIPGPGTKVHVIFNFKGNVLNKD
IRCKDETHLYTLIVRPDNTVEKIDNSQVESGLEDWDPLPKKIKDPDAKPED
WDERAKIDPDTDSKPEDDKPEHIDPDAAKPEDWDEMDGEWEPVTONPEYKGEWK
PRQIDNPDKYTHIPEIDNPEYSPDANIYDSFAVLGLDLWQKSGTIFDNFLITN
DEYAEFEGNETWGTAKAEKMDKQDEBQRLKDEEEDKKKEBEEAEKDEDDDRD
EDEDEBEKEDEEESPGQAKDEL"
BASE COUNT 504 a 453 c 477 g 422 t
ORIGIN

Query Match 64.8%; Score 1368.4; DB 94; Length 1856;
Best Local Similarity 83.1%; Pred. No. 6.4e-244;
Matches 1568; Conservative 0; Mismatches 276; Indels 44; Gaps 9;

QY 74 ttaaaggccgcgcttgcgcgcctcgg-cgcgcctatctctatccgtgcgcttgcg 132
Db 1 TAAAGGCTGTGTGCGCGCGCGCGCTCGCGCGCGCATCTCTTTCGGTGCCTCG 60
QY 133 ctgcgctcctgcgcttgcgcgcctcgcgcgcctcgcgcgcctcgcgcgcctcgc 192
Db 61 CTGTGCTCTCGCTGCGCGCGCGCGCGAGACCTCGCATCTATTTCAGAGCAGTCTTG 120
QY 193 gacgagacggttgactccgcgcctcgcgcgcctcgcgcgcctcgcgcgcctcgc 252
Db 121 GACGGAGATGCTGGACCAACCGCTGGTTCGAATCCAAACATCAAGTCCGATTTGGCAA 180
QY 253 ttgcgttcagtcgcgcgcctcgcgcgcctcgcgcgcctcgcgcgcctcgcgcgc 312
Db 181 TTGTCTCAGTTCGTGCGAAATTTACGGGGACCTGGAGAGGATTAAGGGCTGCAGACA 240
QY 313 agccaggatgcacgctttatctctgcgcgcctcgcgcgcctcgcgcgcctcgc 372
Db 241 AGCCAAGATGCCGATTTTACGCACTGTCCGCCAATTCGAACCTTCAGCAATTAAGGC 300
QY 373 cagcgcgctggttgagtcacgctgaaacatagcagaaacatcgcactgctggcgccg 432
Db 301 CAGACACTGCTGCTACAGTTCACGCTGGAAGCATGACGAGAATATCGACTGTGGGGCGGC 360
QY 433 tatgtgaacttttccctaattgttgaccagacagacatcagcggagactcagaatac 492
Db 361 TACGTGAAGCTGTTCGAGTGGCTGTGGACCAAGAGGACATCATGGAGACTCAGAAAT 420
QY 493 aacatcatgtttgctccgcacatctgtgcccctggcaccacaagaagttcattcttc 552
Db 421 AACATCATGTTTGGTTCGGGACATCTCGGCTCTCGCACCACCAAGGTTTCATGTCATCTT 480
QY 553 aactacaaggccaaagcgtgctgatcaacaaggacatccgttgaagatgatgagttt 612
Db 481 AACTCAAGGGCAAGAAATGCTGTATCAACAAGGATATCCCGGTGAAGGATGATGAATTC 540
QY 613 acacacctgtacactgattgtgcgcgcagacacacacctatgaggtgaagattgacac 672
Db 541 ACACACCTATACACTGATTGTGCGGCGAGACACACCTATGAGGTGAAAATGACAA 600

QY 673 agccaggtgagtcgcgcctccttggaagacgattgggacttctgcccacccaagaata 732
Db 601 AGCCAGGTGGAGTCAGCTCCTTTGGAGGATGATTGGGACTTTCTGCCACCCAGAGATA 660
QY 733 aaggatcctgagtctctcaaacccgggaagactgggtagcggggcccaagatcgtatgcc 792
Db 661 AAGGACCTCTGATGCTCCAAAGCCGGAAGACTTGGGATGAACGAGCCAAAGATTCGATGACCC 720
QY 793 acagactccaagcctgagactggggaagccgagacccagacatactcctaccctgctgaag 852
Db 721 ACAGATTCCAAGCCTGAGGACTGGGACAAGCCAGACACATCCTGACCTGATGCTAAG 780
QY 853 aagcccgagactgggatgaagagatggacggagatgggaaccccccagtgatcagaac 912
Db 781 AAGCCTGAGGACTGGGATGAAGAGATGGATGGAGAGTGGGAACCAACCATGATTCAAAAT 840
QY 913 cctgagtacaagggtgagtgggaagccccggcagatgcacaacccagattacaaggcact 972
Db 841 CCTGAATACAAGGGCGAGTGGAAACCCACCGTCAAAATGACAACCCAGATTACAAGGGTACC 900
QY 973 tggatcccccagaaattgacaaccccgagatttctccgagatcccatctatctgctat 1032
Db 901 TGGATACACCCAGAAATGACAACCCCTGAATACTCCCCCGATGCAAAATATCTATGCTAT 960
QY 1033 gataacttttggcgtgcgtggccctggacctctggcaggtcaagctctgcacacctcttgac 1092
Db 961 GATAGTTTGTCTACTGGCGCTAGATCTCTGCGAGGTCAAGTCCGGGACAACTCTTTGAC 1020
QY 1093 aacttctcatccacaacgtagaggcatacgtctgaggagtttggcaacgagacgtggggc 1152
Db 1021 AATTTCTCATCACCATAATGATAGGCTATGAGGCTATGAGAGGAGTTTGGCAATGAGACGTGGGT 1080
QY 1153 gtaacaaagccagagagaacaaatgaaggacaaacagagacagagagagagccttaag 1212
Db 1081 GTTACCAAGGCTGCAGAGAAGCAGATGAAGGACAAAGCATGAGAGGATGAGAGGAGGCTTAAG 1140
QY 1213 gaggaggaagaagaacaaacgcaagagagagagagagagagagagagagagatgat 1272
Db 1141 GAAGAAGAGAGGACAAAGAAGCGTAAAGAGGAGAGAGAGCTGAGGATTAAGAGGATGAT 1200
QY 1273 gaggacaaagatgaggatgaggagatgaggagagacaaagaggaagatgaggaggaagat 1332
Db 1201 GATGACAGAGATGAAGATGAGGATGAAGAAGATGAGAAGGAGGAAGAT---GAGGAAGAA 1257
QY 1333 gtcccccgcagcccaagacagacgtgttagag-----aggcctgcctccagggcgtgac 1386
Db 1258 TCCCTTGGCCCAAGCAAGGATGAGCTGTAGAGGCCACACACCTGCTTTCAGGGCTGGAC 1317
QY 1387 tgaggcctgagcgtcctgcgcagagcttgcgcgcgcacaaataatgtctctgtgagactc 1446
Db 1318 TGAGGCTGAACAC-CCTGCCCGAGAGCTGGCTGCTCCCAATAAATGCTCTATGAGACTC 1376
QY 1447 ggaacttcatatttttccagcgtggttcgagattggggtgagattgttggttgttccc 1506
Db 1377 AAGAACTTTTCATTTTTCAGGAGGTTTCAGATCTGGGCTAGATTCTGATTTGTGTCC 1436
QY 1507 ctctccactctccccaccctcccccttttttttttttttttttttttaactggtatt 1566
Db 1437 CTCG-----CTCCCCCAATTACCCCCCCCCCTTTTCTTACTGTG-----T 1478
QY 1567 ttatcttggattctcttcagccctcaccctcgttctctcatcttcttctgatacaatctt 1626
Db 1479 TTGTCCTTAATTTCTCTTCAGGCCCTCATCTGGTGTCTCATTTTGTGNAACAATCTTTTC 1538
QY 1627 ttctgtctctgtccct 1686
Db 1539 CTCTCTGCTCCCT 1598
QY 1687 gt---ggagaagccacagcctgagatttctctctctctctctctctctctctctctctct 1743
Db 1599 GTAGAGGAGAGCCCTAGGCTTGGATTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1658
QY 1744 ggcagcagaagggtgtgtctccacccccccagcactgaggaagaagcgggctctctct 1803


```

Db 1659 GGCAGGAAAGGGGTGGTGTTCCTCCCGCCGCTCAGGAGAAATGGGCTCTTCT 1718
Qy 1804 catttcccccctctcttccctgccccagagactggccacttctggtgggcaggt 1863
Db 1719 CAT-----CCCTTTCCTTCCTGCCCCCAGACATGGGCCACTT----GTGGGGCAGC 1766
Qy 1864 ggggtccagattggctcacactgagaatgtaagaactacacaaacaaattcttataatt 1923
Db 1767 CAGTCTAGCACACGCTCACACTGAGAGTGTGAAGAACTACAAACAATTTCTATTAATT 1826
Qy 1924 aaatttgtgtctcaaaaaaataaaaaa 1951
Db 1827 AAGTTTTGTGTCTCAAAAAAATAAAAAA 1854

RESULT 8
RNCALRET RNCALRET 1882 bp mRNA ROD 31-MAR-1995
LOCUS Rat mRNA for calreticulin.
DEFINITION X53363
ACCESSION X53363
VERSION X53363.1 GI:55854
KEYWORDS calreticulin.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE 1 (bases 1 to 1882)
AUTHORS Murthy,K.K.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-1990) Murthy K.K., McGill University, Royal
            Victoria Hospital, 687, Pine Avenue West, Montreal, Quebec H3A 1A1,
            Canada
REFERENCE 2 (bases 1 to 1882)
AUTHORS Murthy,K.K., Banville,D., Srikanth,C.B., Carrier,F., Holmes,C.,
            Bell,A. and Patel,Y.C.
TITLE Structural homology between the rat calreticulin gene product and
            the Oncohercra volvulus antigen Ral-1
JOURNAL Nucleic Acids Res. 18 (16), 4933 (1990)
MEDLINE 90370496
FEATURES             Location/Qualifiers
     source           1..1882
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="brain cortex"
                     /clone_lib="rat brain cdna"
                     /clone="RB 39"
sig_peptide         36..86
                     /product="signal peptide (AA -17 to -1)"
CDS                 36..1286
                     /codon_start=1
                     /product="precursor (AA -17 to 399)"
                     /protein_id="CAA37446.1"
                     /db_xref="GI:55855"
                     /db_xref="SWISS-PROT:P18418"
                     /translation="MLLSVPLLLGLLGLAAADPAIYFKEOFLDGDWNTNRWVSKHS
                     DEGKFFVLSGKFFGDQKDKGLQTSQDAFYALSARPEPFSNKGQTLVQFTVKHEQN
                     IDCGGYVLFPGGLQDKDHGDSSEYNMFPGDTCGTRKRVHVIENYKGNVLINPD
                     IRCKDEFHLYTLVLRPNTYEYKIDNSQVESGSLDDWDFLPPRKIKDPDAKPED
                     WDERAKIDPTDSKPEDDKPEHIPDPAKPEDWDEMDGEWPEPVIQNPYKGEWK
                     PRQIDNPYKGTWIHPEIDNPYSPDANIYAYDSFAVLGLDLWQVRSGLTFDNLFTN
                     DEAYAEFGNETGWVTKAAEKQMKQDEQRLKEBEDKKEEEDKEDDEDDRD
                     EDEDEDEKEDDEDATGQAKDEL"
mat_peptide         87..1283
BASE COUNT          525 a 457 c 484 g 416 t
ORIGIN

Query Match 61.7%; Score 1207.4; DB 95; Length 1882;
Best Local Similarity 81.7%; Pred. No. 1e-231;
Matches 1542; Conservative 0; Mismatches 301; Indels 44; Gaps 11;
```

```
QY 1158 aaaggcagcagagaaacaaatgaagacacaaacaggacgagcagagcgttaagagaga 1217
Db 1085 CAAGGCTGCAGAGAACAGATGAAGACAAAGCAGGATGAGGACGAGCTTAAGGAAGA 1144
QY 1218 ggaagaagacagaaacgcaagagagagagagagagagagagagagagagagagaga 1277
Db 1145 AGAAGAAGACAAAGAGCGTAAGAGAGAAAGAGAGGCGCGAGGATTAAGAGAGATGAGGATGA 1204
QY 1278 caaagatgagatgagagagatgagagagacaaagagagagagagagagagagagatgcc 1337
Db 1205 CAGAGATGAAGATCAAGATGAAGAGATGAGAGAGAGAGAGATGAGAGAG---GATCCAC 1261
QY 1338 cgccaggccaaagagagagctgtagag-----agccctgcctccagggctgagactgag 1391
Db 1262 TGCCCAAGCCAAAGATGAGCTGTAGAGGCCACACCACTTCCTCCAGGCTGGACTGAGG 1321
QY 1392 cctgagcgtcctgcgcagagcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1451
Db 1322 CCTGA--ACACCGGCACAGAGCTGGCTGCTCC--ATAATGTCTCTATGAGACTCAAGAA 1378
QY 1452 cttt---catttttccagctgttgcgagatttgggtgagatttgggttggcttccctc 1509
Db 1379 CTTTTCATTTTTTCCAGCGAGGTTCAAGTCTGGGGTGGATCTCATTTTTTGTTCCTCC 1438
QY 1510 ctccactctcccccacccctccgcctcttttttttttttttttttttttttttttttttttt 1569
Db 1439 TGCTTCCCTCCCAATCCTCCCTCCCAACCCCTTTTCTTTTACTGGTG-----TTTG 1489
QY 1570 tttttgattctcttcagcctcaccctcctcctcctcctcctcctcctcctcctcctcctcct 1629
Db 1490 TCTTTAAATTTTCAGGCTTCAACCTCCTGGCTCTCATTTTTTGA--ATCAACATTTTTTC 1547
QY 1630 ttgcctctgctccctctctctctctctctctctctctctctcctcctcctcctcctcctc 1689
Db 1548 TTCTGTCTCCTTTCTCACTCTTGTGCTACTATCTTCACTTCACTTCACTTCACTTCACTT 1607
QY 1690 gagaagccagagcctgagatttctctctctctctctctctctctctctctctctctctctct 1749
Db 1608 GAAAGCCCTAGGCTGAGATTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1667
QY 1750 agaaggggtgtgtgtc-tccaaaccccccagcactgaggaagcagggggtctctctcattt 1808
Db 1668 AGAAGGGGTGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1721
QY 1809 caccctcctctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1868
Db 1722 -----CCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1770
QY 1869 ccagattggtcacactgagaatgtaagaaactacaaacaaatttctattataaatt 1928
Db 1771 CTAGCACACTCACACTGAGAGTGAAGAACTTACAAACAAATTTCTATTAAATTAAGTT 1830
QY 1929 ttgtgtctcaaaaaaaagagaa 1955
Db 1831 TTGTGTCTTCAAAAAAAGAAAAA 1857

RESULT 9
RATRC3
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (strain:Sprague-Dawley) male testis spermatogenic
cdna to mRNA, clone:RC3.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 1816)
Nakamura,M., Moriya,M., Baba,T., Michikawa,Y., Yamanobe,T.,
Arai,K., Okinaga,S. and Kobayashi,T.
```

```
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

An endoplasmic reticulum protein, calreticulin, is transported into
the acrosome of rat sperm
Exp. Cell Res. 205 (1), 101-110 (1993)
93202172
2 (bases 1 to 1816)
Nakamura,M.
Direct Submission
Submitted (14-NOV-1995) to the DDBJ/EMBL/GenBank databases.
Masahisa Nakamura, Hiroshima University, Faculty of Science,
Laboratory for Amphibian Biology, 1-3-1 Kagamiyama,
Higashi-Hiroshima, Hiroshima 739-8526, Japan
(E-mail:nakamura@ipc.hiroshima-u.ac.jp, Tel:0824-24-7483,
Fax:0824-24-7483)
Location/Qualifiers
1. 1816
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/cell_type="spermatogenic"
/clone="RC3"
/sex="male"
/tissue_type="testis"
16. 1286
/codon_start=1
/product="calreticulin"
/protein_id="BAA11345.1"
/db_xref="GI:1845572"
/translation="MLLSVPLLLGLLAAADPAIYFKQEFLDGDATNRWVSKHKS
DFGRVLSGKGFQDQKDKLQTSQDARFALSARFPFSNKGQTVLVQVTVKHEQN
IDCGGYVKLFPGLDQDMHDSYINFMFGDPCGTGKVVHVFYKNGKVLINKD
IRCKDEFTHLYTLIVPDNTYEVKIDDSQVSESGLEDWDFLPKPKIDPAKPED
WDERAKIDDPDTSKPEDWKPPIIDPAKPEDWDEMDGEWEPVIONPYKGEWK
PROINDPYKGTWHPIDNPEYSPDANIYADSPAVLGDLDMOVKSGTIFDNFLITN
DEATAERGNETWGTAKAEQKMDQDEORLKEEEDKKRKEEEDKEDDEDDED
EDEDEDEKEDEDEDEATQAKDEL"
BASE COUNT 480 a 447 c 474 g 415 t
ORIGIN

Query Match 61.4%; Score 1201.4; DB 95; Length 1816;
Best Local Similarity 82.1%; Pred. No. 1.7e-230;
Matches 1520; Conservative 0; Mismatches 281; Indels 50; Gaps 10;

QY 95 cccctcgccgccatgctctactcagggagcagtttctgacgcgagcgggtggaactccc 214
Db 2 CCCCTCGGCCCGCCATGCTCTCTTGGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 61
QY 155 tcgcgagcctcgctctacttcaaggagcagtttctgacgcgagcgggtggaactccc 214
Db 62 CGCGACACCTGCCATCTATTTCAAAGACAGTCTTGGACGAGATGCTTGACCAACC 121
QY 215 gctgacgaatccaaacacaaagtcagatttggcaaatctgtctcagcttcgccaagt 274
Db 122 GCTGGTCAATCCAAACATAAGTCTGATTTTGGCAATTCGCTCAGTTTGGCAAT 181
QY 275 tctacggtgacagagagaaataaaggttgcacacagcagcagcagcagcagcagcagcag 334
Db 182 TCTACGGGACCGAGAGAGGATTAAGGTTTGCACACACCAAGATGCCGATTTTACG 241
QY 335 ctctcgccagcttctcagccttcagcaacaaaggccagcagcgtggtggtgcagttca 394
Db 242 CGCTGCGCCAGATTCGAACCTTCAGCAACAAGGGCCAGACACTGGTGGTACAGTTCA 301
QY 395 cggtaaacatagcagacacatcgactgtggggcggttatgtgaagctgtttcctaata 454
Db 302 CCGTGAAGCATGAGCAGAAATATCGACTGTGGGGGGGGTACGTACGTGTTTCCGGGTG 361
QY 455 gttgacacagacacatgacgcagcagcagcagcagcagcagcagcagcagcagcagcag 514
Db 362 GCTTGGACGAGACATGCATGGAGACTCAGATATATTAACATCATGTTTGTTCGGACA 421
QY 515 tctgtgcccctggcaccagaaggttcattgtcatcttcaactacaaggcagaagcagcgtgc 574
Db 515 tctgtgcccctggcaccagaaggttcattgtcatcttcaactacaaggcagaagcagcgtgc 574
```

An endoplasmic reticulum protein, calreticulin, is transported into the acrosome of rat sperm
Exp. Cell Res. 205 (1), 101-110 (1993)
93202172
2 (bases 1 to 1816)
Nakamura,M.
Direct Submission
Submitted (14-NOV-1995) to the DDBJ/EMBL/GenBank databases.
Masahisa Nakamura, Hiroshima University, Faculty of Science, Laboratory for Amphibian Biology, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan
(E-mail:nakamura@ipc.hiroshima-u.ac.jp, Tel:0824-24-7483, Fax:0824-24-7483)

Location/Qualifiers
1. 1816
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/cell_type="spermatogenic"
/clone="RC3"
/sex="male"
/tissue_type="testis"
16. 1286
/codon_start=1
/product="calreticulin"
/protein_id="BAA11345.1"
/db_xref="GI:1845572"
/translation="MLLSVPLLLGLLAAADPAIYFKQEFLDGDATNRWVSKHKS
DFGRVLSGKGFQDQKDKLQTSQDARFALSARFPFSNKGQTVLVQVTVKHEQN
IDCGGYVKLFPGLDQDMHDSYINFMFGDPCGTGKVVHVFYKNGKVLINKD
IRCKDEFTHLYTLIVPDNTYEVKIDDSQVSESGLEDWDFLPKPKIDPAKPED
WDERAKIDDPDTSKPEDWKPPIIDPAKPEDWDEMDGEWEPVIONPYKGEWK
PROINDPYKGTWHPIDNPEYSPDANIYADSPAVLGDLDMOVKSGTIFDNFLITN
DEATAERGNETWGTAKAEQKMDQDEORLKEEEDKKRKEEEDKEDDEDDED
EDEDEDEKEDEDEDEATQAKDEL"

BASE COUNT 480 a 447 c 474 g 415 t
ORIGIN

Query Match 61.4%; Score 1201.4; DB 95; Length 1816;
Best Local Similarity 82.1%; Pred. No. 1.7e-230;
Matches 1520; Conservative 0; Mismatches 281; Indels 50; Gaps 10;

QY 95 cccctcgccgccatgctctactcagggagcagtttctgacgcgagcgggtggaactccc 214
Db 2 CCCCTCGGCCCGCCATGCTCTCTTGGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 61
QY 155 tcgcgagcctcgctctacttcaaggagcagtttctgacgcgagcgggtggaactccc 214
Db 62 CGCGACACCTGCCATCTATTTCAAAGACAGTCTTGGACGAGATGCTTGACCAACC 121
QY 215 gctgacgaatccaaacacaaagtcagatttggcaaatctgtctcagcttcgccaagt 274
Db 122 GCTGGTCAATCCAAACATAAGTCTGATTTTGGCAATTCGCTCAGTTTGGCAAT 181
QY 275 tctacggtgacagagagaaataaaggttgcacacagcagcagcagcagcagcagcagcag 334
Db 182 TCTACGGGACCGAGAGAGGATTAAGGTTTGCACACACCAAGATGCCGATTTTACG 241
QY 335 ctctcgccagcttctcagccttcagcaacaaaggccagcagcgtggtggtgcagttca 394
Db 242 CGCTGCGCCAGATTCGAACCTTCAGCAACAAGGGCCAGACACTGGTGGTACAGTTCA 301
QY 395 cggtaaacatagcagacacatcgactgtggggcggttatgtgaagctgtttcctaata 454
Db 302 CCGTGAAGCATGAGCAGAAATATCGACTGTGGGGGGGGTACGTACGTGTTTCCGGGTG 361
QY 455 gttgacacagacacatgacgcagcagcagcagcagcagcagcagcagcagcagcagcag 514
Db 362 GCTTGGACGAGACATGCATGGAGACTCAGATATATTAACATCATGTTTGTTCGGACA 421
QY 515 tctgtgcccctggcaccagaaggttcattgtcatcttcaactacaaggcagaagcagcgtgc 574
Db 515 tctgtgcccctggcaccagaaggttcattgtcatcttcaactacaaggcagaagcagcgtgc 574

```
Db 422 TCTCGGTCTCTGGCACCAGAAGGTTTCATCTCTTTAACTACAAGGGCAGAAGCTGC 481
Qy 575 tgatacaaggacatcgttgcaagatgatgatttacacacactgacacactgattg 634
Db 482 TGATCAACAAGGATATCGGTGTAAAGATGATGAATTCACACATCTATACACGCTGATTG 541
Qy 635 tgcggccagacaacacatctagagtggaattgacaacagccaggtgagtcgggctcct 694
Db 542 TGCGCCAGACAACACCTACGAGGTGAATAATTGACAACAGCCAGGTGCGGCTCCT 601
Qy 695 tggagacgattgggaattcctgcaccccaagaagataaaggatcctgatgcttcaaac 754
Db 602 TGGAGGATGATTGGGATTTCTGCCGCCAAGAAGATTAAGGATCCTGACCTGCCAAGC 661
Qy 755 cggagactggatgagcggcgaagatcatgatccacagactccaagcctgagact 814
Db 662 CAGAAGACTGGGATGAACGAGCCAAAGATTGATGACCCACAGATTCGAAGCCTGAGACT 721
Qy 815 gggacaagccgagcatatcctcctgacctgatgctaagaagccgagactgggatgaag 874
Db 722 GGGACAAGCCAGAGACATCCTGACCTGATGCTAAGAAGCCTGAGACTGGGACGAAG 781
Qy 875 agatggagcagagatgggaacccccagtgattcaagaacctgatacaagggtgagtgga 934
Db 782 AGATGGATGGAGATGGGAACCCACAGTGATTCAAAATCTCTGAATACAAGGGCGAATGGA 841
Qy 935 agccccgcagatcgacaacccagattacaaggcacttggtatccaccagaaattgaca 994
Db 842 AGCCACGTCAAAATTGACAACCCAGATTACAGGGTACTGGATACACCAGAGATTGACA 901
Qy 995 acccagagtattcccgatcccgatctatgctctatgataaatttggcgtgctggcc 1054
Db 902 ATCTGTAATACTCCCGGATGCGAATATCTATGCTATGATAGTTTCTGCTACTGGCT 961
Qy 1055 tggacctctggcagtgcaagtctggaccattcttggacacttctcattaccacacagtg 1114
Db 962 TAGACCTCTGGCAGTCAAGTCTGGCAATTTTGTGACAACTTCTCATCAACCAATGATG 1021
Qy 1115 agccatagctgagagatttggcaacagagcgtggcgctaaacaagcagcagagaac 1174
Db 1022 AGGCTATGACAGGAGTGTGGCAATGAGACCTGGGTGTACCAAGGCTGCAGAGAGC 1081
Qy 1175 aaatgaggcaaacagcagcagcagcagagcgttgaaggaggaggaagaagacagaac 1234
Db 1082 AGATGAAGGACACAGCAGGATGAGGACAGAGCCTTAAGGAAGAAGAGACAGAAGC 1141
Qy 1235 gcaagaggagagagcagcagagcagagcagcagcagcagcagcagcagcagcagcagc 1294
Db 1142 GTAAAGAGGAGAGGAGGCGCAGGATAAAGAGGATGAGGATGACAGAGATGAAGTGAAG 1201
Qy 1295 aggatgaggagcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1354
Db 1202 ATGAAGAGGATGAGAAGGAAGAAGATGAGGAG---GATGCCACTGGCCCAAGGATG 1258
Qy 1355 agctgtagag-----agggcctcctcagcgtgagcgtgagcgtgagcgtcctgcgc 1408
Db 1259 AGCTGTAGAGGCCACACACCTGCTCAGCGCTGGACTGAGGCTGAGCCTGAACAC-CCCGCCA 1317
Qy 1409 cagagcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1466
Db 1318 CAGAGCTGGCTGCTCCCAATAATCTCTATGAGACTCAAGAAGCTTTTCATTTTTTTTC 1377
Qy 1467 agcgtgtcggatttggggtgatttgggttgggttccctcctcctcctcctcctcctcctc 1526
Db 1378 AGCGAGGTTAGGCTGGGGTGGATCTGATTTTGTGTC-----CCTGCTCCTCCCAATCC 1432
Qy 1527 ccctcccccgcctcttttttttttttttttaaaactgggtatttttttttttttttttttt 1586
Db 1433 TCCCCACCCCTTTTCTTTTACTGGTGTGCTTTAATT-----CTTCA 1478
Qy 1587 gccctcacccttggtctcattcttcttgatcaacatttttttttttttttttttttttttt 1646
Db 1479 GCCCTACCTCTGGCTCTCATTTTGAATCAACATTTTCTTTTCTTTTCTTTTCTTTTCTCT 1537
```

```
Qy 1647 tctcatcttagctccctccacacacacacacacacacacacacacacacacacacacacac 1706
Db 1538 CATCTCTGGTCACTATCTCTCAACTTAGGAACAGGTATGAGGAAGACCTAGGCTTG 1597
Qy 1707 agattcatctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1765
Db 1598 AGATTTTCATCTGCTCTCTCTTCTGAATCTCAGAGAGGGTAGGAGAGGGGCTGCTATCT 1657
Qy 1766 tcaaacccccacacacacacacacacacacacacacacacacacacacacacacacacac 1825
Db 1658 TCCTTCCCTCCAGCAGCTGAGGAAGATGGGGCTCTTC-----CCCTTCTTCC 1704
Qy 1826 cctccccccagactgggcccactctggtggcgagtggtccccagattggctcacact 1885
Db 1705 CTTCCTCCCTCAGAGCTGGCCATTT---GTGGGCGACGCTTCTAGCACAGCTCACACT 1760
Qy 1886 gagaatgtaagaactacaacacacacacacacacacacacacacacacacacacacacacac 1936
Db 1761 GAGAGTGAAGACTACAACAAAATTTCTATAAATTAACTTTTGTGTCT 1811
```

RESULT 10

```
RABCBP
LOCUS RABCBP 1865 bp mRNA MAM 27-APR-1993
DEFINITION Rabbit calcium binding protein (calreticulin) mRNA, complete cds.
ACCESSION J05138
VERSION J05138.1 GI:164858
KEYWORDS calcium-binding protein; calreticulin.
SOURCE Rabbit (neonatal) skeletal muscle sarcoplasmic reticulum, cDNA to mRNA, clones pTZ[18R,19R].
ORGANISM Oryctolagus cuniculus
REFERENCE 1 (bases 1 to 1865)
AUTHORS Flegel,L., Burns,K., MacLennan,D.H., Reithmeier,R.A.F. and Michalak,M.
TITLE Molecular cloning of the high affinity calcium-binding protein (calreticulin) of skeletal muscle sarcoplasmic reticulum
JOURNAL J. Biol. Chem. 264, 21522-21528 (1989)
MEDLINE 90094320
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted by M.Michalak, 29-SEP-1989.
```

FEATURES

```
source
1..1865
/organism="Oryctolagus cuniculus"
/db_xref="taxon:9986"
<1..1865
/note="CBP mRNA"
61..111
sig_peptide
61..1317
CDS
61..1317
/note="calreticulin signal peptide"
/note="calreticulin precursor"
/codon_start=1
/protein_id="AAA31188.1"
/db_xref="GI:164859"
/translation="MLLPVPLLLGLGLAAEPVYFKEOPLDGDGWTWIERWIESKHKS
DFKGFVLSGKFFYDQEKDGLQTSODARFVALSAREPESNKGOPVLVQVTKHEON
IDCGGTVKLFPAQDQKDMHDSSEYINMGPDTCGPTKKYHVFNFKKNVLINKD
IRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGLEDWDFLPPKKIKDPDASKED
WDRKATIDDPDTSKPEWDRKPEHIDPDAKKPEWDEEMGEWPEVIQNPYKGEWK
PROINDPDYKGTWHPIDNPEYSPDANIYAVDSFVGLDLMQVKSQGTIDNPLITN
DEAVAEFGNETGWTKTAOKMKDKODEQRLEKEEEKKREKEEAEDEEDKDDK
EDEDEDEEDKDEEEEAQAQAKDEL"
mat_peptide
112..1314
/note="calreticulin"
BASE COUNT 427 a 543 c 579 g 316 t
ORIGIN
```

```
Query Match 56.1%; Score 1098.6; DB 7; Length 1865;
Best local Similarity 78.0%; Pred. No. 6.3e-210;
Matches 1489; Conservative 0; Mismatches 349; Indels 71; Gaps 11;
```

QY 54 gccgtgccggagggttcgttttaaaagg---ccgcgctgttcgcgcctccctcgccc---gcc 108
Db 1 GCCGCTGCCGGAGGATCGTTTAAAGGGCCCGCGAGCGCGCCCTCGGCCCTGCC 60
QY 109 atgctgctatccgtgctgtctgctgcctcctcctgcgtgcgtgcgtccgcagcctgccc 168
Db 61 ATGCTGCTCCCTGTGCGCTGTGCTGCTGCGGCTGTGCTGCGGCTGCGGCCCGCGCCGCTC 120
QY 169 gtctacttcaaggagcagttcttgagcagagacgggtggaacttccogctggatcgaatcc 228
Db 121 GTCCTACTTCAAGAGAGCAGTTTCTGGACGGAGATGGGTGACCGAGCGCTGGATTCGAATCC 180
QY 229 aaacacaaagtcaagattttgcaaaatcgtctcagttccgcgcaagtctacaggtgacgag 288
Db 181 AAACACAAGTCGATTTTGGCAAAATTCGTCTCAGTTTCGGGCAAGTTTCACGGCGATCAG 240
QY 289 gagaagataaaagggtttgcagacaagccaggatgcacgctttttatgctctgtcgcgcagt 348
Db 241 GAGAAAGATAAAGGGCTGCAGACCAAGCCAGGACGCCCGCTTCTACGCCCTGTGCGGCCGA 300
QY 349 ttccagcctttcagcaaaagccagacgcgtggtgagcttcacgttcaacacatgag 408
Db 301 TTCGAGCCCTTCAGCAACAAAGGCCAGCCAGCCACTGGTGGTGCAGTTTCAACCAACAGAG 360
QY 409 cagaacatcgactgtggtggcggtctatgtgaagctgttttccctaatagtttgaccagaca 468
Db 361 CAGAAATTTGACTGCGGGGGCGCTACGTGAAGCTGTTTCCGGCCGGCGCTTGACCAAGAG 420
QY 469 gacatcacggagactcagaatacaacatcatgtttggtcccgacatctgtggtgcctggc 528
Db 421 GACATGCACGGGACCTCTCAGTACAAACATCATGTTTGGTCCCTGACATCTGTGGCCCGGC 480
QY 529 accaagaaggttcactatcttcaactacaagggaagaaacgctgctatcaacaaggac 588
Db 481 ACCAAGAAGGTTTCACTCATCTCTCAACTCAAGGGCAAGAACGTGCTGATCAACAAGGAC 540
QY 589 atccgtgtcaagatgatgatttacacacctgtacacactgattgtgcgcccagacaac 648
Db 541 ATCCGTTGCACAGGACACGAGTTTACACACCTGTACACCTGATGTCGCGCGCGGACAC 600
QY 649 acctatggttgaagattgacaacgcgggtggagtcggctcccttggaaagacattgg 708
Db 601 ACCTATGAGTGAAGATTGACAACACCCAGGTGGAGTCCGGTCCCTGGAGGATGACTGG 660
QY 709 gacttctcccaagaaagataaaggatcctgattcttcaaaaacccggaagactgggat 768
Db 661 GACTTTCCTACCCCCCAAGAAAGATAAAGGACCCAGATGCCCTCGAAGCCTGAAGACTGGGAC 720
QY 769 gagcgggcccaagatcgatgatccacagactccaagcctgagactggagacagcccgag 828
Db 721 GAGCGGGCCCAAGTGCACGACCCACCGGACTCCAGCCCGAGACTGGGACCAAGCCCGAG 780
QY 829 catatccctgacccctgatgctaagaagcccgaggactgggtgaagagatgacagagagag 888
Db 781 CACATCCCCGACCCGACGCGAAGAACGCCGAAGACTGGGACCAAGAAATGACGCGGAGAG 840
QY 889 tgggaacccccagtgattcagaacccctgaagtacacaagggtgagtgaagcccccgagatcc 948
Db 841 TGGGAGCCGCCCGGTGATTTCAGAACCCCGAGTACAAGGGTGAAGTGAAGCCCGGCAGATTC 900
QY 949 gacaacccagatataaagggtcacttgatcccccagaaattgacaacccccgagtattct 1008
Db 901 GACACCCCGATTACAAAGGACCTTGGATCCACCCCGGAATTCACACACCCCGAGTACTCG 960
QY 1009 cccgatcccgatctatctatgataaactttggcgtgctggcgtggactctggcag 1068
Db 961 CCGCAGCGCTTAACATCTATGCCCTACGACAGCTTTCGCGTGTGCGGCTTGACCTCTGGCAG 1020
QY 1069 gtccagctcggcaccatctttgacaacttcctcatcaccacaagatgagcattacgctgag 1128
Db 1021 GTCAAGTGGGGACCATCTCTCACAACATCTCTCATCATCAACCAAGATGAGCGGTACCGCAG 1080

QY 1129 gaggttggcaacgagagactggggcggtacaaaggcgagagagaaacaaatgaaggacaaa 1188
Db 1081 GAGTTTGGCAACGAGAGACTGGGGCGTCAACAAGACGCCGAGAGAGAGAGAGAGAGAG 1140
QY 1189 caggacagagagagagaggttaaggaggggaagaagacaaagaaacgcaaaagggaggag 1248
Db 1141 CAGGACAGGAGCAGCGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1249 ggggc---agagacaagagagatgatgagacaagatgaggaagatgagagatgagag 1305
Db 1201 GAGGCCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1306 gacaagaggaagatgagaggaagatgtcccgccagggcccaagggacagctgtagaga 1365
Db 1261 GACAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1366 ggcctgcctccagggtgcgtgactgagggcctgcgtgcctgcctgcctgcctgcctgcct 1425
Db 1321 GGCTTCC-----AGGCCGCCGAGGCCCTGAGCGCTC-----GGCGCGAGAGCCA 1364
QY 1426 aataatgtctctgtgagactcgagaacttt---cattttttccagagctgggttcgattt 1482
Db 1365 AATAATGTCTCTATGAGACTCGAACTTCTCCCCACCTCTCCCGAGGAGGTTTCAGATTC 1424
QY 1483 ggggtgatttgggtttgttccctcctccactctcccccacccctcccgccctttt 1542
Db 1425 GGGTGGGTATTTTGGTTTGTTC-----CCCTCCCCCA 1461
QY 1543 tttttatttttaaaactgggtatttttatcttcttcttcttcttcttcttcttcttcttct 1602
Db 1462 TTTTTTTTTTTTTACTGGCGTTTGAATCTTGAATCTCC-----CCGTAAGCCCTC 1516
QY 1603 ctcatcttcttgatcaacatcttttttcttctctgt-----cccttctctct 1652
Db 1517 CCGCTCTCTCAATCAGCATCTTCCCTTCTCCCTGGGTGCGCTGCTTCTTCTCCCTCAG 1576
QY 1653 ctctlagctccctccacacctggggcggtggtgtggagaagccacagggcctgagattt 1712
Db 1577 ATGTGTTAGGCCACCTCTGTAAACCCAGAGAGTGGGAGCGGTGTGGAGACGAGCCCC 1636
QY 1713 catctgctctcttcttgagccagagagagggcagcagagaggggtg-----tgctctc 1768
Db 1637 CCAGCTCGAGCTCCCGGGAGGGGTGGGGGTGGGGGAGATGGCAGGAGAAAGAGGG 1696
QY 1769 aacccccagactgaggaagaacgggctcttcttcttcttcttcttcttcttcttcttct 1828
Db 1697 GACGCTCCGGCCACCCCTCAGCTGACCTGACGAGGAGGAGTAGGCTCCCGCCCTTG 1756
QY 1829 gccccagactgggcccactt---ctgggtggggcagtggtggtccagattggtccactg 1886
Db 1757 TCCCCCAGACCGGCCACTTGTGGTGGGGCAGTGGGTTCCAGCTCGGCTCAGACTG 1816
QY 1887 agaattgaagaactacaaacaaatttctattataaatttgaatttggctc 1935
Db 1817 CGAATGTAAGAACTTCAAAACAAATTTCTATTAAATTAAGTTTGTGTC 1865
RESULT 11
RNSPCALR 1539 bp mRNA ROD 21-JUL-1995
LOCUS R.norvegicus (Sprague Dawley) calreticulin mRNA.
DEFINITION X79327
ACCESSION X79327 1 GI:488840
VERSION X79327 1
KEYWORDS calreticulin.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1539)
AUTHORS Sonnichsen,B., Fullekrug,J., Nguyen Van,P., Diekmann,W.,
Robinson,D.G. and Mieskes,G.
TITLE Retention and retrieval: both mechanisms cooperate to maintain

calreticulin in the endoplasmic reticulum
J. Cell. Sci. 107 (Pt 10), 2705-2717 (1994)
95181573
REFERENCE 2 (bases 1 to 1539)
AUTHORS Soennichsen, B.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-1994) B. Soennichsen, Abt. Klinische Biochemie,
Universitaet Goettingen, Robert-Koch-Strasse 40, 37075 Goettingen,
FRG

FEATURES		Location/Qualifiers	
source	1..1539	/organism="Rattus norvegicus"	
		/strain="Sprague Dawley"	
		/db_xref="taxon:10116"	
		/tissue_type="liver"	
		/clone_lib="lambda zapII cDNA"	
		/clone="CR-BS"	
	14..1264	/gene="calreticulin"	
	14..1264	/gene="calreticulin"	
		/codon_start=1	
		/product="calreticulin"	
gene		/protein_id="CAA55890.1"	
		/db_xref="GI:488841"	
		/db_xref="SWISS-PROT:P18418"	
		/translation="MCLSVPLLLGLGLAAADPAITYKFEQFLDGDAWTNRWVESKHS DRGKFLVSGKYVQDQKDKGDTQSDARYFALSARPEFNSKQTLVVOFTVKHBEON IDCGGYVKLFPGGLQDQKDGSEYNIMFGDPICGPTKVKHVIYNYGKGNVLINKD IRCKDEFHLTVLIVRPNTVEKIDNSOVESGLEDDEWDFLPKPKIKDPPAAKPED WDERAKIDPDTSKPEDDWKPEHIDPDAKPKEDWDEMDGWEPEVIGNPEYKGEWK PQIDNPDKYGTWIIHPEIDNPXPDPANIYAVDSFVGLGLDLQVKSAGTIFDNFLITN DEAYAEFNETGWTKAAEKQMKDKQBEQRLKEEEDKKRKEEEAEDEKDEDDDD EDEDEFEKNEEEDATGQAQDEL"	
	BASE COUNT	421 a	369 c 409 g 340 t
	ORIGIN		
	Query Match	54.3%;	Score 1063.4; DB 95; Length 1539;
	Best Local Similarity	83.4%;	Pred. No. 7e-203;
	Matches 1289; Conservative	0;	Mismatches 226; Indels 30; Gaps 6;
	Qy	96	ccccctggccgcctgctctactcgttcgctgtgctctcgtcggcctcctcgtcgcctgcccgt 155
Db	1	CCCTGGCCGCCGATGCTCTTTCGTTGGCGCTCCTGCTTGGCTCCTCGGCTCGGCTG 60	
Qy	156	cgcgcagcctgcgcgtctactcctcaaggagcaggtttctggacggagacgggtggacgtcccg 215	
Db	61	CGCAGACCCCTGCCATCTATTTTCAAGAGCAGTCTTGGACGGAGATGCCCTGGACCAACCG 120	
Qy	216	ctggatcgatccaaacacacagtcagatttggcaaatctgtctcagttccggcaagtt 275	
Db	121	CTGGGTGCAATCCAAACATTAAGTCTGATTTTGGCAAAATTCGTCTCAGTTCTGGCAAAAT 180	
Qy	276	ctacggtgacgagagaaagataaaggcttgcagacacgacgagtcacgctttttatgc 335	
Db	181	CTACGGGACACGAGAGAGATTAAGGTTGCAGACAAGCAAGATCCCGCATTTTACGC 240	
Qy	336	tctgtcggccagtttgcagccttccagcaaaaggccagacgctggtgcagttcac 395	
Db	241	GCTGTCCGCAGATTCGAACCCCTTCAGCAACAAAGGGCCAGACACTGGTGGTACAGTTTAC 300	
Qy	396	ggtgaacatgacgagacacatcgttggggcggtatgtgaagctgtttccctaag 455	
Db	301	CGTGAAGCATGAGCAGAAATATCGACTGTGGGGCGGCTACGTGAAGCTGTTTCCGGGTGG 360	
Qy	456	tttggacacagacatgcacgagactcagatacaacatcatgtttgttcccgacat 515	
Db	361	CTTGGACCAAGACATGATCGAGACTCAGATATATACATATATGTTGGTCCGGACAT 420	
Qy	516	cttggccctggccacgaaggttcatgtctatcttcaactcaaaaggccgaagcgtgct 575	
Db	421	CTGGGTCTCTGGCCACCAAGAGGTTTCTATGTCATCTTTAACTACAAGGGCAAGACGTGCT 480	

Qy	576	gatacaacagagacatccgtttgcaaggatgatgagtttacacacctgtacacactgtattgt 635
Db	481	GATCAACAAGSATATCCGGTGAAGGATGATGAATTTACACATCTATACACGCTGATTGT 540
Qy	636	gcggccagacaacacctatgaggtgaagattgacaacagccagggtgagtcgactcctt 695
Db	541	CGGCCCCAGACAACACCTTACGAGGTGAAATTTGACAACAGCCAGGTGGAGTGGGCTCTT 600
Qy	696	ggaagacgattggacttctgccaccacaagaataaagatcctgactgcttcaaaacc 755
Db	601	GGAGGATGATGGGACTTTCGCCGCCCAAGAAGATTTAAGGATCTCAGCTGCCAGGCC 660
Qy	756	ggaagactggatgagcggccaaagatcgatcccccacagactccaagcctgaggactg 815
Db	661	AGAAGACTGGATGAACGAGCCAAAGATTGATGACCCACACAGATTTCCAAGCCTGAGGACTG 720
Qy	816	ggacaagcccgagcatatcctgacctgatgctaaagaagcccgagactggatgaaga 875
Db	721	GGACAAGCCAGACACATCCCTGACCTGATGCTTAAGAAGCCTGAGGACTGGGACGAAGA 780
Qy	876	gatgacgagagatggaaaccccccagtgattcagaacccctgagtacaaggtgagtgaa 935
Db	781	GATGGATGGAGAGTGGGAACCCACAGTGATTCAAAATCTCTGAATACAAAGGCGGAATGGA 840
Qy	936	gccccgcagatcgacacccacagattacaagggcacttggatccaccaccagaaattgcaa 995
Db	841	GCCACGTCAAATTGACAACCCAGATTACAAGGTTACCTGGATACACCACAGAGATTGACAA 900
Qy	996	ccccgaatatctccagatcccgatctctatgctatgataacttttggcgtgctgggcct 1055
Db	901	TCCTGAATATCTCCCCCGATCGAATATCTATGCTCTATGCTATGCTTCTGCTGCTGCTT 960
Qy	1056	ggacctctgcagtcgaactctgcacatctttgacaacttctcatcaccacacagatga 1115
Db	961	AGACCTCTGGCAGGTCAAGTCTGGCACAAATTTTGTGACAACTTCTCTATCACCATGATGA 1020
Qy	1116	ggcatacgtcgtgaggtttggcaacgagcgtggggcggttaacaaagcagcagagaaca 1175
Db	1021	GGCCTATGCAGAGAGTTCGCAATGAGACCTGGGTGTCAACCAAGCTGCAGAGAAGCA 1080
Qy	1176	aatgaagacaacacagacgagcagagcgtttaaaggaggaggaagaagaagaagaacg 1235
Db	1081	GATGAAGGACAAGCAGGATGAGGAGCAGAGGCTTTAAGGAAGAAGAAGAAGAAGACG 1140
Qy	1236	caagagagagagagcagacagacaaaggagatgatgagacaaagatgaggtgagga 1295
Db	1141	TAAAGAGGAAGAGAGGCCGAGGATAAGAGGATGAGGATGACAGAGATGAAGATGAAGA 1200
Qy	1296	ggatgaggaggaacaaggaggaagatgaggagaagatgtccccggccaggccaaaggacga 1355
Db	1201	TGAAGAGGATGAGAGGAAGAGAGATGAGGAG---GATGCCACTGGCCCAAGCAAGGATGA 1257
Qy	1356	gctgtagag-----aggcctgcctccaggcgtggaactgaggcgtgagcgtcctgcgcg 1409
Db	1258	GCTGTAGAGGCCACACACACCTGCCTCCAGGCTGGACTGAGGCTTGA---ACACCCGCCAC 1315
Qy	1410	agagcttcgcgcgccaataatgtctctgtgagactcgagaac-tttcattttttttccag 1468
Db	1316	AGAGCTGGCTGCTCC-AATAATGCTCTATGAGACTCAAGAACTTTTTCATTTTTTTTCAG 1374
Qy	1469	gctggttcgagattgggtgagatttggtttgcctcctcctcactctccccaccacc 1528
Db	1375	GCAGGTTACGTCGTGGGTGGATTCGATTTTTCCTCCCTGCCCTCCCATCTCCCCAC 1434
Qy	1529	ctcccccccttttttttttttttttaactgggttatatttatcttttattctctcagc 1588
Db	1435	CCCCCTTTTTTTTTTACTTGGTGT-----GTCTTTAATTTCTTCAGC 1477
Qy	1589	cctcaacccctggctctctattcttcttgatcaacaacatctttttcttgc 1633
Db	1478	CCTCACCTCTGGCTCTCATTTTTTGAATCAACATTTTTTTCTTTTC 1522

```

RESULT 12
CH19HHR23
LOCUS
DEFINITION
  CH19HHR23 110096 bp DNA PRI 01-APR-1997
  Homo sapiens DNA from chromosome 19p13.2 cosmids R31240, R30272 and
  R28349 containing the EXLF, GCDH, CRTG, and RAD23A genes, genomic
  sequence.
ACCESSION
  AD000092
VERSION
  AD000092.1 GI:1905905
SOURCE
  human.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 110096)
  Lamerdin J., McCready P., Stillwagen S., Ramirez M. and Carrano A.
  Characterization by genomic sequence analysis of a gene-rich 111 kb
  region of 19p13.2 containing the human DNA repair gene, RAD23A
  Unpublished
JOURNAL
  Lamerdin J.E.
REFERENCE
  2 (bases 1 to 110096)
  Direct Submission
  Submitted (08-NOV-1996) J.E. Lamerdin, Human Genome Center,
  Lawrence Livermore National Laboratory, 7000 East Ave, Livermore,
  CA, USA, 94551 jane@acgt.llnl.gov ow@ornak.llnl.gov
COMMENT
  GSDB:S:985657.
  map=19p13.2.
FEATURES
  Location/Qualifiers
    1..110096
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /chromosome="19"
     /cell_lines="5HL2-B"
     /cell_type="fibroblast"
     /clone_lib="LL19NC03 R chromosome 19-specific cosmid
     library"
     /map="19p13.2"
     /note="constructed at LLNL from flow-sorted chromosomes
     from hybrid 5HL2-B, which carries chromosome 19 as its
     only human chromosome"
  8..86
   repeat_region
     /note="repeat match = HSAL09846; putative"
     /rpt_family="Alu"
  24..407
   repeat_region
     /note="repeat match = HSAL02504; putative"
     /rpt_family="Alu"
  132..407
   repeat_region
     /note="repeat match = HSAL06234; putative"
     /rpt_family="Alu"
  289..437
   repeat_region
     /note="repeat match = HSAL05358; putative"
     /rpt_family="Alu"
  complement(474..648)
   misc_feature
     /note="predicted exon, gail2exons_human_1.3; frame=0,
     reverse strand, quality=excellent; putative"
  499..537
   misc_feature
     /note="ss region (gap in top strand); putative"
     /complement(518..651)
     /note="repeat match = L1MB7; putative"
     /rpt_family="LINE"
  complement(648..949)
   repeat_region
     /note="repeat match = HSAL02217; putative"
     /rpt_family="Alu"
  complement(953..1105)
   repeat_region
     /note="repeat match = L1MB7; putative"
     /rpt_family="LINE"
  1068..1179
   misc_feature
     /note="ss region (gap in bottom strand); putative"
     /complement(1465..1576)
     /note="repeat match = HSAL00829; putative"
     /rpt_family="Alu"
  1489..11769
   repeat_region
     /note="repeat match = ALU; putative"

```

```

/rpt_family="Alu"
1634..1806
  /note="repeat match = HSAL00499; putative"
  /rpt_family="Alu"
1843..1995
  /note="repeat match = HSAL02158; putative"
  /rpt_family="Alu"
1879..2001
  /note="repeat match = HSAL11628; putative"
  /rpt_family="Alu"
  complement(2081..2169)
  /note="predicted exon, gail2exons_human_1.3; frame=2,
  reverse strand, quality=good; putative"
  complement(2833..3137)
  /note="repeat match = HSAL03055; putative"
  /rpt_family="Alu"
  complement(3617..4230)
  /note="repeat match = HSAL02794; putative"
  /rpt_family="Alu"
  complement(4200..4239)
  /note="repeat match = HSAL01548; putative"
  /rpt_family="Alu"
  4274..4557
  /note="repeat match = HSAL06523; putative"
  /rpt_family="Alu"
  4282..4489
  /note="repeat match = HSAL02768; putative"
  /rpt_family="Alu"
  complement(4492..4666)
  /note="predicted exon, gail2exons_human_1.3; frame=0,
  reverse strand, quality=excellent; putative"
  4702..4969
  /note="predicted exon, gail2exons_human_1.3; frame=2,
  forward strand, quality=excellent; putative"
  4887..4970
  /note="similarity: pir||A54602; microtubule-associated
  serine/threonine protein kinase MAST205; putative"
  join(<4890..4970>5060..5267,11345..11483,11582..11714,
  11852..12017,12253..12354,12516..12638,13189..13298,
  14009..14187,14285..14512,15235..15368,15489..15830,
  17359..17481,17571..17707,19859..20099,20158..21405)
  /note="coding region constructed from xgrail 1.3
  predictions, EST and SWISSPROT matches; comment for
  location 17571-17707: BLASTN similarity to H14890:
  ym25g07.rl Homo sapiens cDNA (bases 198-325); (5.9e-57);
  user-supplied translation (frame +1):
  GRSSKAKPPGGENDFTIKLISNGAYGVLYVRHRTQRFAMKKINKONLILRNQIQ
  QAFVERDILTEA
  ENPFVGMFCSEFTRRHLCMVMEYVEGGDCATLLKNICALPVEMARMYFAETVLALEY
  LHNYSIVHRDLK
  PDNLITSMGHIKLTDFGLSKWGLSLTNYEGHIEKDAREFLDKQVCGTPEYIAPE
  VILRQGGKPKVD
  WNWAGIILYEFVLGVCVPPFGTPELFGQVTSDDILWPEGDEALPTEAQLLISSLLQT
  NPLVRLGAGGAF
  EVKHGSFRLDLDWTGLLRQKAEIFPHLESEDDTSYFDRSDRYHHVNSYDEDDTTEE
  PVEIRQFSSCP
  RFKSVYSMEQLSQHEPKTPVAAAGSSKREPSTGKPEKVACKREGGLGLILREKTWR
  GGSPEIKRFSAS
  GGSPEIKRFSAS
  EASFLEGEASPLGARRRFSALLPEPSRFSAPQDEDEARLRPRPRSSDDPAGSLDARA
  AGDSEASPRATNDLVRARRHQMSGDVAVEKRPRTGGKVIKASATASALSAGPMRG
  SATALGGGRSRY
  REMSIFVLVLSDHGCVSVVDPHGSSPLASPMSPSSLSNPSRSSPSRDYSPAVSGL
  RSPITIORSGK
  YGFTLRIRVYMGDTDVYSVHHIVHVEEGGPAQAGLCAGDLITHVNGEPVHGMVHP
  EVVELILKSGNK
  VAVTTPFENTSIIRIGPARRSSYKAKMARNRKPSAKEGQESKKRSSLFRKITQSNL
  LHTSKSLSLNR
  SLSSDSILPGSPHTGLPARSPTHSYRSTPDSAYLIGTCTCAGTEQRCQVAVLSSPAS
  STPNSPASSASH
  HIRPSTLHLSJPSKLUHQVRSARCKSAGNIPLSPLAHTPSPPTQASPPPLPGHTVSSHT
  TQSPFAPKLHSSP
  PVRPRPKSABPPRSPILLKRVQSAEKLGLASLADKKGALRKHSLEVGHVDFRKFPHGE

```

LALHLSAESDGE
TPVVEGLGAPRQVAVRRLGRQESPLSLGADPLLPEGASRPPVSSKEKESPGGAECTP
PRATTFGGRTLE
RDVGTCTRHQSVQTFEDGTGGMARAVAKAALSPVOEHETGRSSSGEAGTPLVPIVVEPA
RPGKAKVVPOL
GADSKGLOEPAPLAPSVEAPRGRERKRVLEVVEERTLSGPRSKPASPKLSPPEQPPS
LAPAKCASPSA
VTPVPASLIGSTKQVGLTSCRPAAEAVPPAGLTKKGVSSPAPPGP; comment
for location 5060-5267; BLASTN similarity to U02313:Mus
MAST205 protein kinase mRNA (bases 1510- 1720); (4.0e-50);
comment for location 20158-21405; BLASTN similarity to
D61343: Human fetal brain cDNA 5'-end (bases 1- 188);
(1.1e-32); comment for location 12516-12638; BLASTN
similarity to U02313:Mus MAST205 protein kinase mRNA
(bases 2253- 2382); (1.0e-18); comment for location
13189-13198; BLASTN similarity to U02313:Mus MAST205
protein kinase mRNA (bases 2388- 2492); (4.0e-12); comment
for location 11582-11714; BLASTN similarity to T99935:
ye2d02.r1 Homo sapiens cDNA (bases 192- 294); (1.5e-42);
comment for location 11345-11483; BLASTN similarity to
T99935: ye72d02.r1 Homo sapiens cDNA (bases 53- 195);
(1.5e-30) and U02313:Mus MAST205 protein kinase mRNA
(bases 1716- 1858); (1.2e-28); comment for location
11852-12017; BLASTN similarity to U02313:Mus MAST205
protein kinase mRNA (bases 1989- 2157); (5.2e-37); comment
for location 17359-17481; BLASTN similarity to U02313:Mus
MAST205 protein kinase mRNA (bases 3344- 3467); (1.8e-19)
and H14890: ym25g07.r1 Homo sapiens cDNA (bases 75- 197);
(2.6e-42); comment for location 15489-15830; BLASTN
similarity to U02313:Mus MAST205 protein kinase mRNA
(bases 3117- 3344); (4.2e-40) and H14890: ym25g07.r1 Homo
sapiens cDNA (bases 2-74); (4.2e-20); putative
/codon_start=1
/product="hypothetical human serine-threonine protein
kinase R31240.1"
/protein_id="AA051171.1"
/db_xref="GI:1905906"
/translation="GRSSKAKPPGENDFTIKLSNGAYGAVLVLRHRTQRFAFK
KINKQNLIRNOIQOAFVERDILTFAPNPVGMFCFETRHLCMMVMEVGGDCAT
LKNIGALPEVMARMAFAETVLALEYLHNGYIVHDLKPDNLITSGHILKIDFGLS
KMGLSLTTNLEGHIEKDAREFLDKQVCQTPPEYIAPEVILRGYKGPVDWAMGILL
YELFVGCPFGDTPELFQVISDDILMPEDGAUPTAEQLISLLQTNPLVRLGA
GFAPEVHQSFFDLQKQAEFIHLESEDDTSYFEDSDRYHHVNSYDEDDT
TEEPVEIRFQSCDFSKVYSMEQLSQHEPKTPVAAAGSSKRPSTPGPEKAVG
KREGGLGLTURETKWGGSPDIKRFASAESEFLGEASPLGARRRFSALLESRSFA
POQDEEARLRPRSSDPAGSIDARAPKEETQEGTSSAGDSEASPRATNDVLFR
ARQOMSGDVAVKRSRTGGKVIKASATALSAGPMRGCSATLGGGRYREMSIF
LVLSIDHGCVSDVDPHGSSPLSPSPSSNPSSSDSPSDYSPAVSGLRSPITI
QRSKKYGTFLRAIRVYMGTDVYSVHHIYVHVEEGPAQEAAGLDLITHNGEPV
HGMVHPVEVLIKSGNKVAVTTTPFENTISIRIGPARRSYKAKMARNRKRPASKEG
ESKRSFLPKITKQNLHITSLSLNSLSSDSLPGSPHGLPARSPHSYST
PDAYLIGTCTCAGTEORGVANLSSPASSTPNSSASASHIRPSTLHGLSPKLHR
QYRSACKSNGNIPLELAHTPPTQASPPPLGHHVGSSTTQSPFAKHSPPVVR
PRPKSAEPPRLKRVQSAEKIGLASADKGLRHSLEVGHPFRKDFHGLALH
SLAESDGETPVEGLAPQVAVRRLGRQESPLSLGADPLLPEGASRPPVSSKEKES
GGAECTPPRATTPGGTTLRDVGTCTRHQSVQTFEDGTGGMARAVAKAALSPVOEHETG
RRSSSGEAGTPLVPIVVEPARPKAVVPQPLGADSKGLOEPAPLAPSVEAPRGRER

Query Match 39.7%; Score 778; DB 91; Length 110096;
Best Local Similarity 90.9%; Pred. No. 9.6e-146;
Matches 874; Conservative 0; Mismatches 5; Indels 83; Gaps 1;

Qy 1059 cctctggcagctcaagctctggcaccattcttgacaaactctcatcaccacgatgagc 1118
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90088 CATCTCCAGTCAAGTCTGGCACCATCTTTGACAACTTCTCTATCACCACGATGAGGC 90147
Qy 1119 atacgtgagagtttggcaacagacgtggcgctaaaca----- 1159
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90148 ATACGCTGAGGAGTTTGGCAACGAGACGTGGGCGGCTAACAAAGGTGAGGCCCTGCTCG 90207
Qy 1160 ----- 1159
Db 90208 TCCTGATGTCTGGGGCGGCGACGAGGCTGGCAGGGGGCAAGGCCCTCTGCTGCTGCG 90267

Qy 1160 ----aggcagcagagaaaaaatgaagacaaaacagagcagagagagagcttaagag 1215
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90268 CTGCAGGCAGCAGAGAAACAAATGAAGACAAACAGGACGAGGACGAGGCTTAAGGAG 90327
Qy 1216 gagaaagaacacagaacacgaaagagagagagagagagagagagagagatgatag 1275
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90328 GAGGAAGAAGACAAAGAAACGCAAGAGAGGAGGAGGAGGACAAAGAGGATGATGAG 90387
Qy 1276 gacaaagatgagatgagagagatgagagagagagagagagagagagagagatgtc 1335
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90388 CACAAAGATGAGGATGAGGAGGATGAGGAGGACAAAGAGGAGGATGAGGAGGAAGATGTC 90447
Qy 1336 cccggcaggcccaaggacagctgtagagagagcctgcctccaggcctgagctgagcctg 1395
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90448 CCGGCCAGGCCAACAGGACGAGCTGTAGAGAGGCTGCCTCCAGGCGCTGAGCTGAGGCCCTG 90507
Qy 1396 agcctcctgcgcagagacttgcgcgcgcaaaataatgtctctgtgagactcgagaaacttt 1455
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90508 AGCGTCTCGTCGCGAGAGCTTGC CGCGCCAAATAATGTCTCTGTGAGACTCGAGAACTTT 90567
Qy 1456 catttttccagactggttcggaatttgggtggaatttgggttttgggttccctccctccac 1515
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90568 CATTTTTTCCAGGCTGGTTTCGGATTTGGGGTGGATTTGGTTTGTTCCTCCCTCCTCCAC 90627
Qy 1516 tctcccccacccccctccccgcctttttttatttttttaaaactgggtattttatctttg 1575
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90628 TCTCCCCACCCCTCCCGCCCTTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 90687
Qy 1576 attctcttcagcctcaccctgggtctctctctctctctctctctctctctctctctgct 1635
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90688 ATTCTCTTTCAGCCTCACCCTGGTTCTCATCTTCTTCTGATCAACATCTTTTCTTGCT 90747
Qy 1636 ctgtccctctctctctctctctctctctctccctcccaactggggggcagtggttgagaag 1695
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90748 CTGTCCCCCTTCTCATCTCTTAGTCTCCCTCCCAACCTCGGGGGGCGAGTGGTGTGAGAAG 90807
Qy 1696 ccacagcctgagatttct 1755
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90808 CCACAGGCTCAGATTTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 90867
Qy 1756 ggggtgtgtctccacccccccagcactgaggaagacggggtctctctctctctctctctct 1815
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90868 GGGTGGTGTCTCCAAACCCCGCAGCACTCAGGAAGAACGGGCTCTTCTCATTTTCAACCCCT 90927
Qy 1816 cctttctcctctgccccagactgggccaacttctggtggggcagtggttccagatt 1875
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90928 CCTTTCTCCCTGCCCCAGACTTGGGCCACTTCTGGGTGGGCGAGTGGGTGCCAGATT 90987
Qy 1876 ggtccacactgagaatgtaagaactacaaacaaatttctataattaaattttgtgttc 1935
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 90988 GGCTCACACTGAGAAATGTAAAGACTACAAACAAATTTCTATTAATTAATTTTGTGTC 91047
Qy 1936 tc 1937
||
Db 91048 TC 91049

RESULT 13
BOVCALRET 1542 bp mRNA MAM 21-OCT-1993
LOCUS Bovine calreticulin (brain isoform) mRNA, complete cds.
DEFINITION L13462
ACCESSION L13462
VERSION L13462.1 GI:348693
KEYWORDS calcium binding; calcium binding protein; calreticulin.
SOURCE Bos taurus (library: lambda gt11) brain cDNA to mRNA.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (sites)
AUTHORS Liu, N., Fine, R.E. and Johnson, R.J.
TITLE Comparison of cDNAs for bovine brain coding for two isoforms of

JOURNAL Biochim. Biophys. Acta 1202 (1), 70-76 (1993)
MEDLINE 9385184
REFERENCE 2 (sites)
AUTHORS Liu,N.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1993) via Mac Floppy by Genbank
FEATURES
Location/Qualifiers
source
1..1542
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="brain"
/tissue_lib="lambda gt11"
223..324
sig_peptide
CDS
223..1488
/note="brain isoform"
/codon_start=1
/product="calreticulin"
/protein_id="AAC37307.1"
/db_xref="GI:348694"
/translation="MCLNHLFSLVLSIVLLPHEVFIICLHHIVTFLRETVFFSEOF
LFLDLKYKASKLSSIREALSMSKVGIIENFCSEISFLQESIKSHGRITLVGCSPPGH
EQNIDCGGYVNPFPAGLDQDMGDSYINMFGPDICGPGTKKVVIFNTKGRNVL
INKDIRKDETHLYTLIVRPNTYEVKIDNSQVSGSLDDWDFLPKKIKDPDAA
KPEDWDRAKIDDPDSDKPDHPIDPDAPKPEDEMDGEMGEWEPPLIQNDPEYK
GEWKPRQIDNPKGIWIHPEDINPEYSPDSNIYAYENFAVLGLDLWVKSTIFDNF
LITNDVAAEPEFNETGWVTKAAEKQMDKQDEQRLHEEBEKKKEEAEKDDDE
DKDEDEDEDEEEEDAAAQAKDEL"
BASE COUNT 429 a 319 c 403 g 391 t
ORIGIN

Query Match 39.3%; Score 770; DB 7; Length 1542;
Best Local Similarity 86.2%; Pred No. 3.8e-144;
Matches 864; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY 404 atgagacaatcgactgtggggcggtatgtgaagctgtttcctaataagtttgacc 463
Db 530 AAGACAGAACATCGACTGTGGGGCGGCTATGTGAACGTGTTTCCAGCTGGTTGGATC 589
QY 464 agacagacatgacgagactagaatacaatacatgtttgttccgcacatctgtggcc 523
Db 590 AGACAGACATGCGACGAGACTCCGATACACATAATGTTGGCCCGGACATCTGTGGCC 649
QY 524 ctggcaccagaaggttcgtcatcttcaactacaagcgcaagaacgctgctgataca 583
Db 650 CTGGCACCAAGAGGTTCTATGTCATCTTCAACTACAGCGCAGAGATGCTGTATCAACA 709
QY 584 aggacatccgttgcaggatgatgtattacacacctgtacacactgattgtgctggccag 643
Db 710 AGGATATCCGGTGAAGGATGATGAATTCACCCACCTGTACACGCTGATTGTGCGGCCA 769
QY 644 acaacacatgaggtgagattgacaacagccagggtgagtcggctcctctggaagacg 703
Db 770 ATAATACCTATGAGTGAAGATGACACAGCCAGGTGAGTGCAGGCTCTTGTGGAGGACG 829
QY 704 attggacttctgccaccagaagataaagatcctatctcctcaaacaggagact 763
Db 830 ATTGGATTCTTTCGCCACCAAGAGATGAAGATCTCTGATCCGCTAAGCCCTGAAGACT 889
QY 764 gggatgagcggccaaagatcgatgattccacagactccaagcctgaggactgggacaagc 823
Db 890 GGGACGATCGCCCAAGATCGATGACCTTACAGACTCCAAAGCTGAGGATTGGGACAAGC 949
QY 824 ccgagacatacctgacccctgattgataagaccccgagagactggagatgagacg 883
Db 950 CTGACACATCTCTGACCCCTGATGCTAAGAAACCTGAGGACTGGGATGAAGATGGAGC 1009
QY 884 gagatgggaaccccgatgattcaaaccttgatcaaacaggtgagtggaagcccgccg 943
Db 1010 GAGAGTGGGAACCACTCTTATTCAGAACCCAGAGTACAAAGGGGAGTGGAAACCCAGC 1069
QY 944 agatcgacaacccagattacaagggcactgttgattccaccagaaattgacaaccccgagt 1003

|||||
Db 1070 AGATGCAAAACCCAGAGTCAAGGCGATTGGATCCATCCAGAGATTGACAACCCCTGAGT 1129
QY 1004 attctcccgatccagtatctatgcctatgataactttggcgtgctggcctgagacctct 1063
Db 1130 ATTCCCTTGACAGCAACATCTATGCTTATGAAACTTCGCTGTCTAGGCTTGATCTTT 1189
QY 1064 ggcaggtcaagtctggcaccatctttgacaacttctcatcaccacgagatgagcacaag 1123
Db 1190 GGCAGGTCAAGTCTGGCACCATTCTTGACAACTTCTCTCATCACCACGATGAAGCGTATG 1249
QY 1124 ctgagagatttggcaacgagacgtgggggtaacaaaggcagcagagaaacaaatgaag 1183
Db 1250 CTGAGGAGTTTGGCAACGAGACGCTGGGTGTTTACAAAGGCAGCAGAAACAAATGAAGG 1309
QY 1184 acaaacagagacgagagcagagccttaagagagaggaagaagaacacgaagaagag 1243
Db 1310 ACAAGCAAGATGAAGCAGCAGGCTACATGAGGAGGAGGAGAGAAAGGCAAGGAGG 1369
QY 1244 agagagagcagaggaagagagatgatgagacaaagatgagatgagagagatgag 1303
Db 1370 AGGAAGGCGACAGAAAGATGATGACGAGACAAAGGATGAGGATGAGGAGATGAAGATG 1429
QY 1304 aggacaagagaggaagatgagaggaagatgtccccgcaggcccaaggagcagctgtaga 1363
Db 1430 AGAAGCAAGAGGAGGAGGAAGAAGATGCTGCCGCTGCGCAGGCCAAGGATGAGCTGAGG 1489
QY 1364 g---aggctgctccaggcgtgagctgagcctgagcgcctc 1402
Db 1490 GCCACCTCTGCTTCCAGGGCTGGACTGAGGCTGAGTCGTC 1531

RESULT 14
XLOCUS XLCALR 1291 bp mRNA VRT 18-JAN-1993
DEFINITION X.laevis mRNA for calreticulin (clone 3).
ACCESSION X67597
VERSION X67597.1 GI:64608
KEYWORDS calreticulin.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 1291)
AUTHORS Treves,S.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1992) S. Treves, General Pathology, Via Borsari 46, 44100 Ferrara, ITALY
REFERENCE 2 (bases 1 to 1291)
AUTHORS Treves,S., Zorzato,F. and Pozzan,T.
TITLE Identification of calreticulin isoform in the CNS
JOURNAL Biochem. J. In press
FEATURES Location/Qualifiers
source
1..1291
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/tissue_type="CNS"
/clone_lib="lambda gt11"
/clone="3"
sig_peptide <1..36
CDS <1..1236
/codon_start=1
/product="calreticulin"
/protein_id="CAA47866.1"
/db_xref="GI:64609"
/translation="LVPLLAGLCAIEPAVYFKEEFTDGDGTQHWESKHKTGYGK
KLSAGFYDSEKQKGLTQSDARFYAMSSRFSFNKQDTLVVQVFSVKHQNIDCGG
GYKLFPAALEQTEHMESEYNIMFGPDICGPTTKKHVIFQYKKKNQIKNDIRKCD
DSFTHLYTLIVRPNTYEVKIDNSQVSGSLDDWDFLPKKIKDPKPEWDERP
KIDDPKPKPEWDERPPIPDPAVKPEDWDEMDGEWEPVQNPYKGEWKPRQID

polyA_site
BASE COUNT 665 a 560 c 589 g 580 t
ORIGIN
DDDDDEEEKEEEEESEAPQKDEL"
2394

Query Match	36.2%	Score 709.2;	DB 8;	Length 2394;
Best Local Similarity	74.9%	Pred. No. 5.7e-132;		
Matches 888; Conservative	0;	Mismatches 298;	Indels 0;	Gaps 0;

Qy	1108	aacgatgagcatacgcctgaggagtttgcaacagagacgtgggcgctaacaaaggcagca	1167
Db	1015	GACGATGAGAAGTTTTCGACGAGGAACATGCCATTAGACATGGGGAGTCACAAAGGAAGGA	1074
Qy	1168	gagaacaacatgaagacacaaacagacgagggacagaggccttaaggagagaggaaagaac	1227
Db	1075	GAGAAAAGATGAAGGAACAGCAGGATGAAGAGGAACGCCAAGAACACGAGGAGGAAGAG	1134
Qy	1228	aagaaacgcacaaaggagggagggcagaggacaaaggaggaTgtaggagcacaagatgag	1287
Db	1135	AAGAAGAGGAACAGCAGGAGCCTGCTGAAGAGGCTGAGGATGATGACGACGATGATGAT	1194
Qy	1288	gatgaggaggatgaggaggacaaaggagggaagatgaggagggaagatcy	1333
Db	1195	GATCACGACGAGGAAGAAAAGAGAGAGGAAGGAAGGAGGATGAGGAGG	1240

Search completed: October 28, 2001, 18:04:03
Job time: 4492 sec

polyA_site	DDDDDEEKEKEEEDDEESEAPOKDEL*
BASE COUNT	2394
ORIGIN	665 a 560 c 589 g 580 t
Query Match	36.2%; Score 709.2; DB 8; Length 2394;
Best Local Similarity	74.9%; Pred. No. 5.7e-132;
Matches 888; Conservative	0; Mismatches 298; Indels 0; Gaps
Qy	148 ctggccgtccgcagctgcgcgtctacttcaaggagcagtttctgagcgagagcgggtgg 207
Db	55 CTGTGCTCTGCTAAACCGCGTCTACTTCCRAAGAGAAATCGGGATGGAGATGCCCTGG 114
Qy	208 acttccgcgtggatcgaatcccaacacagtcagattttggcaaatggtctctcagttcc 267
Db	115 AAAGAACGCTGGATACAATCTAAGCACAAAGTCAGACTACGGCCAGCTTAAAGCTGAGTGC 174
Qy	268 ggcagttctacggtgacgaggagaagataaagtttgcagacaagccagagatgcacgc 327
Db	175 GGCAAGTTTACGGAGATGAAGAGAAGACAAGAGTCTCCAGACCAGCAAGATGCCAAG 234
Qy	328 ttttatgctctgcgcgcagtttccagctttcagcaacaagaagccagacgctagtggtg 387
Db	235 TTCTATGCCACCTCTGCTGGATTCCCTGTTCTTCCAAACAAAGACAAGCCCTCTTGTAAGT 294
Qy	388 cagttcaagctgaaacatgacgagaacatcgactgtggggcgctatgtaagctgttt 447
Db	295 CAGTTCTACGTGAAGCACGACGACAGACATGACTGTGGCGGTGGATGACGTGAAGCTTTTC 354
Qy	448 cctaagtagtttgaccagacagacatgcacggagactcagaatacaacatcatgtttggt 507
Db	355 CCAAGCACTCTAGACGAGACTGACATGTCATGGAGAGTCAGAATACAACATCATGTTGGT 414
Qy	508 ccgcgacatctgtggccctggccaccaagaaggttcagtgcattctcaacatacaaggccaag 567
Db	415 CTTGACATTTTCGGGGCACCCACCAAGAAGTGCACGTCACTCTTCACTACAAGGGGAAG 474
Qy	568 aacgtctgatcaacaagacatccgtgtcaaggatgatgatttacaacactgtacaca 627
Db	475 AACTTGCAGATCAACAAAGATATCGCATCCAAAGCCGATGTTACAGTCATCTCTACACC 534
Qy	628 ctgattgtgcgcgcagacaacactatgaggttgaagattgacaacagccaggtggagtcc 687
Db	535 TTGATTGTAGCCCTGTATACACCTACGAGGTAAAAATTGATAACAGCAAGGTGTAATCT 594
Qy	688 ggctcttggaaacgattgggacttcctgcaccaccaagaagataaaggatactctgatgct 747
Db	595 GGCAACCTGGAAAGACGACTGGGACTTCCCTTCCGCCCAAGAAGTTAAGGACCCCGAGGCC 654
Qy	748 tcaaaaccggaagactgggatgagcgggccaagatcgatgattcccaagactccaagcct 807
Db	655 AAGAAGCCCGATGACTGGGATGAGCGTCCAGATTTGATGATCTCTAGGACAAGAAACCA 714
Qy	808 gaggactggacaagcccgagcatatccctgacctgatgcttaagaagcccgagactgg 867
Db	715 GAGGACTGGACAAGCCCAACAATATCCCTGACCAGATGAGTGAACACAGAGATGG 774
Qy	868 gatgaagatggacggagtgaggaaacccccagtgattcagaacccctgagtacaaaggtt 927
Db	775 GATGAGGAGATGGATGGAGTGGGAGCCACCCAGTCATCAAAATCCAGAGTACAAGGGA 834
Qy	928 gagtgaagccccggcagatcgacaacccagattacaaggcgacttggatccaccacgaa 987
Db	835 GAGTGGAAACCCAGCCAGATTGACAAACCTGACTACAAGGAAAGTGGGTGCACCCAGAG 894
Qy	988 attgacaaccccgagtattctccgactccagtatctctgctctatgataaacttttggcgtg 1047
Db	895 ATCGACAACCCAGAAATATACCCAGATCCTACTCTGTGTACTCCTATGACAGACTTTGGAGCT 954
Qy	1048 ctgggcctggacctgtggcaggtcgaagttcgaccactctttgacaacttctcatcacc 1107
Db	955 TTGGTCTTGTATCTGTGGCAGGTGAATCTGCACCATCTTTGACAACTCTTCGATTACA 1014

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 16:50:51 ; Search time 139.57 Seconds
(without alignments)
8808.708 Million cell updates/sec

Title: US-09-807-148-7
Perfect score: 1958
Sequence: 1 gaatccgcatcttaac.....aaaaaaaaaagaattc 1958

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1958	100.0	1958	21	AAA09346 Human MBP-calretic
2	1837.2	93.8	1890	10	AAN91471 1.9 kb Ro (RO/SSA)
3	1251	63.9	1251	21	AAA09347 Human MBP-calretic
c 4	481.2	24.6	1209	17	AAT39519 Flea calreticulin
5	481.2	24.6	1209	17	AAT39517 Flea calreticulin
6	481.2	24.6	1589	17	AAT39516 Flea calreticulin
c 7	481.2	24.6	1589	17	AAT39518 Flea calreticulin
8	438.8	22.4	1076	12	AAQ11987 Insert from lambda
9	369.8	18.9	385	21	AAC02428 Human secreted pro
10	366	18.7	366	20	AAK27251 Human cClqR bindin
11	366	18.7	1248	21	AAC55204 Human secreted pro

12	363.4	18.6	1514	22	AAF29741	Castor bean calret
13	360	18.4	1513	22	AAF29755	Castor bean calret
14	340.4	17.4	1532	21	AAC40452	Arabidopsis thalia
15	338.8	17.3	1530	21	AAC49247	Arabidopsis thalia
16	337.2	17.2	352	19	AAV56984	Human calreticulin
17	333.6	17.0	1335	21	AAC42868	Arabidopsis thalia
18	317.4	16.2	329	16	AAT20550	Human gene signatu
19	311.6	15.9	366	20	AAK27252	Mouse cClqR bindin
20	302	15.4	366	20	AAK27253	Rat cClqR binding
21	209	10.7	212	21	AAC78267	Human cancer assoc
22	184	9.4	442	20	AAV89143	EST clone CE120.
c 23	148.2	7.6	257	17	AAF39520	Flea calreticulin
24	147	7.5	147	22	AAF68366	Human lung tumour
25	145	7.4	257	17	AAT39515	Flea calreticulin
26	122	6.2	1372	22	AAC89718	Human cytoskeleton
27	108.6	5.5	3191	22	AAK29742	Castor bean calret
28	107	5.5	463	21	AAC48096	Zea mays DNA fragm
29	98.4	5.0	470	21	AAC46669	Zea mays DNA fragm
30	97.8	5.0	3489	21	AAK30290	Kaposi's sarcoma-a
c 31	97.8	5.0	32207	20	AAV73805	KSHV LUR DNA (nucl
c 32	97.8	5.0	137507	19	AAV19941	KSHV long unique c
c 33	95.4	4.9	51259	18	AAK83007	Partial mouse WRN
34	92	4.7	550	21	AAK39388	Cat flea hindgut a
35	91	4.6	1995	21	AAC48757	Arabidopsis thalia
36	91	4.6	1998	21	AAC41110	Arabidopsis thalia
c 37	91	4.6	16442	18	AAK83006	Partial mouse WRN
38	88.2	4.5	363	12	AAO11801	N-terminal sequenc
39	87.6	4.5	1281	21	AAK298333	A. thaliana gene 1
40	84.2	4.3	1392	21	AAK52187	Arabidopsis thalia
41	84.2	4.3	49999	20	AAK23891	Murine LOBO genomi
42	84.2	4.3	49999	20	AAK23896	Murine LOBO genomi
43	83.4	4.3	3100	16	AAK84723	Calnexin DNA seque
44	83.4	4.3	4000	18	AAK91902	Mannose-1-phosphat
45	81.6	4.2	465	21	AAC36294	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAA09346
ID AAA09346 standard; DNA; 1958 BP.
XX
AC AAA09346;
XX
XX
DT 10-AUG-2000 (first entry)
XX
XX Human MBP-calreticulin coding sequence.

XX MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
XX endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
KW cytosolic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic; ss.
OS Homo sapiens.

Key	Location/Qualifiers
FT CDS	109..1362
FT	/*tag= a
FT sig_peptide	109..159
FT	/*tag= b
FT mat_peptide	160..1360
FT	/*tag= c

XX WO200020577-A1.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.

XX 06-OCT-1998; 98US-0103438.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.


```
QY 409 cagaacatcgactgtggggcgctatgtgaagctgtttcctaataagtttggaccagaca 468
Db 301 cagaacatcgactgtggggcgctatgtgaagctgtttcctaataagtttggaccagaca 360
QY 469 gacatgcacggagactcagaataacacatcatgttttggtcccgacatctgtgcccctggc 528
Db 361 gacatgcacggagactcagaataacacatcatgttttggtcccgacatctgtgcccctggc 420
QY 529 accaagaagttcatgtcatcttcaactacaagggcgaagaacgtcgtgatcaacaaggac 588
Db 421 accaagaagttcatgtcatcttcaactacaagggcgaagaacgtcgtgatcaacaaggac 480
QY 589 atccgttgaaggatgatgatgttttacacacctgtacacactgattgtgcggccagacaac 648
Db 481 atccgttgaaggatgatgatgttttacacacctgtacacactgattgtgcggccagacaac 540
QY 649 acctatgaggtgaagattgacaacaccaggtgagtcggctccttgggaagacgattgg 708
Db 541 acctatgaggtgaagattgacaacaccaggtgagtcggctccttgggaagacgattgg 600
QY 709 gacttctgcaccccaagaagataaaggatcctgactgcttcaaaaacgcgaagactggat 768
Db 601 gacttctgcaccccaagaagataaaggatcctgactgcttcaaaaacgcgaagactggat 660
QY 769 gagcgggccaagatcgatgccacagactccaagcctgagactgggacaagcccag 828
Db 661 gagcgggccaagatcgatgccacagactccaagcctgagactgggacaagcccag 720
QY 829 catatccctgacctgatgctaagaagcccgagactgggataagagatgagcagagag 888
Db 721 catatccctgacctgatgctaagaagcccgagactgggataagagatgagcagagag 780
QY 889 tgggaacccccagtgattcagaacctgagtacaagggtgagtggaagccccgcgagatc 948
Db 781 tgggaacccccagtgattcagaacctgagtacaagggtgagtggaagccccgcgagatc 840
QY 949 gaaacccagattacaagggaacttgatccaccgaaattgacaacccccgattct 1008
Db 841 gaaacccagattacaagggaacttgatccaccgaaattgacaacccccgattct 900
QY 1009 ccgataccagtatctatgctatgataactttggcgtgctgggctggacctctggcag 1068
Db 901 ccgataccagtatctatgctatgataactttggcgtgctgggctggacctctggcag 960
QY 1069 gtcaagcttgcaccatcttgacacttctcctcaccacacgatgagcgcactgag 1128
Db 961 gtcaagcttgcaccatcttgacacttctcctcaccacacgatgagcgcactgag 1020
QY 1129 gattttggcaacgagacgtggggcgtaacaaaggcgagagaaacaaatgaaggacaaa 1188
Db 1021 gattttggcaacgagacgtggggcgtaacaaaggcgagagaaacaaatgaaggacaaa 1080
QY 1189 caggacgaggagcagaggcttaaggaggaggaagaagacaaagcaagcgaagggag 1248
Db 1081 caggacgaggagcagaggcttaaggaggaggaagaagacaaagcaagcgaagggag 1140
QY 1249 gaggcagaggacaaaggagatgatgaggacaaagatgaggatgaggagatgaggaggac 1308
Db 1141 gaggcagaggacaaaggagatgatgaggacaaagatgaggatgaggagatgaggaggac 1200
QY 1309 aaggaggaagatgaggaggaagatgtcccgccaggccaagggacgagctg 1359
Db 1201 aaggaggaagatgaggaggaagatgtcccgccaggccaagggacgagctg 1251

RESULT 4
AAT39519/c
ID AAT39519 standard; cDNA; 1209 BP.
XX
AC AAT39519;
XX
DT 12-DEC-1996 (first entry)
XX
```

```
DE Flea calreticulin cDNA clone nCtCall209 (complement).
XX
KW Calreticulin; flea; haematophagous insect; allergic dermatitis;
KW vaccine; therapy; nCtCall209; ss.
XX
OS Ctenocephalides felis.
XX
PN WO9628469-A1.
XX
PD 19-SEP-1996.
XX
PF 08-MAR-1996; 96WO-US03133.
XX
PR 09-MAR-1995; 95US-0401509.
XX
PA (HESK-) HESKA CORP.
XX
PI Rushlow KE, Stiegler GL;
XX
XX WPI; 1996-442861/44.
XX
PT Haematophagous insect calreticulin protein - used to reduce insect
PT infestation and desensitise patients to allergic dermatitis
XX
PS Claim 16; Page 72; 86pp; English.
XX
CC The antisense strand (AAT39519) of nucleic acid nCtCall209 (AAT39517)
CC hybridises under stringent conditions to a flea calreticulin gene.
CC nCtCall209 codes for a flea calreticulin protein (AAW04171) that can
CC be used to alter the blood feeding behaviour of haematophagous
CC insects or to desensitise an animal to allergic dermatitis. The
CC antisense strand, when administered to an animal, reduces
CC calreticulin activity in insects feeding on the animal, thereby
CC reducing the insect burden on the animal and in the environment
CC of the animal.
XX
SQ Sequence 1209 BP; 292 A; 280 C; 212 G; 425 T; 0 other;

Query Match 24.6%; Score 481.2; DB 17; Length 1209;
Best Local Similarity 66.0%; Pred. No. 6.3e-101;
Matches 712; Conservative 0; Mismatches 363; Indels 3; Gaps 1;

QY 240 agattttggcaaatctgtctcagttccggcaagttctacggtgacgagagaagataa 299
Db 1078 AGAATTCGGCAAAATTCGTGCACACTGCCGGAAGATTCTATAACGATGCCGAAGCAGACAA 1019
QY 300 aggtttgcagacaagccaggatgcacgccttttatgtctctgtcggccagtttcgagccttt 359
Db 1018 AGGTTTGCAAAACAGGTCAGATGCTAGGTTCTACGCTCTATCTCATAGTTCAAAACCTTT 959
QY 360 cagcaacaaggccagacgctgtgtggtgcagttccacggtgaacacatgacgagacaatcga 419
Db 958 CTCAATAAAGACAAAGACATTAGTTGTACAAATTTTCGTTTAAACATGAACAAAAACATTGA 899
QY 420 ctgtggggcggtatgtgaagctgtttcctaataagtttgaccagacacacatgcacgg 479
Db 898 CTGTGGAGGTGGTACTTGAAGGGTTTCGAATTCAGTGTGAATCAAAAGGACATGCGATGG 839
QY 480 agactcgaatacaacatcatgtttgtcccgcacatctgtggcctgtggcccaagaaggt 539
Db 838 GGAAGTCCCTATGAATTTATTTGTGCTCTGACATTTGTGACCCAGGAACATAAGAGGT 779
QY 540 tcatgtcatcttcaactacaagggaagaacgctgctgatcaacaagacatcccttgcac 599
Db 778 TCACGTAATCTTCAGCTACAAGGGTAAAAATGTTTGTATCAATAAAGGATATCCGCTGCAA 719
QY 600 ggaatgatgtttacacacctgtacacactgattgtgcggccagacacacccatagagt 659
Db 718 AGATGATGTCTATACTCATGTATACACTTTGGTTGTTTAAAGCCCGATTAATACCTATAGGT 659
QY 660 gaagattgacaacagccaggtgagtcggctccttggaagacgattgggacttccctgcc 719
```

Db 658 GTTGATTGATGATGAGAGTTGGAAGTGGTAACTTTGGAAGATGACTGGGAATTCCTAGC 599
 QY 720 acccaagaagataaagatactctgatctgtctcaaaaccggaagactggatgagcgggccaa 779
 Db 598 CCCCAAGAAAATCAAGGATCCAGAGCTAAAAAACCCAGCAGATGGGATGAGCGTGCAC 539
 QY 780 gatcgatgacccacagactccaagcctgagactggagctggagcccgagacatatccctga 839
 Db 538 TATCTCTGACCTTAATGACACCAAACTGAAGATTGGGACAGCCAGACACATCTCTGA 479
 QY 840 cctgatgtcgaagacccgagactggatggaagatggagcgagtgggaaaccccc 899
 Db 478 TCCTGATGCTACCAACCTGATGACTGGATGAGAGATGGATGGATGGGAACCTGC 419
 QY 900 agtgatcgaaccctgagtacaaaggtgagtgaagcccccgagatgacgaacaccaga 959
 Db 418 TATGATTGACAACCTTGAATATAAGGGAGATGGGCACCAAAAACAGATTGACAATCCTGA 359
 QY 960 ttacaaggcacttgatccaccggaattgacaacccccgagattctcccgatccccag 1019
 Db 358 CTACAAGGTGCTGGGTTACCCCTGAAATGATAATCCAGAAATATGTTCTCTGATACTCA 299
 QY 1020 tatctatgctctatgataaactttggcgtgctggccctggaccttgccaggtcaagtctgg 1079
 Db 298 ACTTTACAAACGTGATGACATTTGTGCCATTTGTTAGATTTATGCAAGTAAAGCTGG 239
 QY 1080 caccatcttgacaactctcctacacacgatgaggcatacgtgaggtttggcaa 1139
 Db 238 AACAATATTCGACAATATTTAATCACACATGATGTTGATTATGCAAGAAATAG--C 182
 QY 1140 cgagcgtggcggttaacaaaggcagcagagaaacaaatgaaggcaaacacgagcagga 1199
 Db 181 AGAAGGTGTTAAATCTTACCAGGAGGAGAGAAAGAAAATGAAGATGCTCAAGATGAAGA 122
 QY 1200 gcagagcgttaaggagggaagacaaagaaacgcgaagaggagggagagcagagga 1259
 Db 121 AGAAGGAAGCCAGGAGCCGCAACAAAGAAAGAAATGACACAGATGCTGATGAAGA 62
 QY 1260 caaggaggatgatgaggacaaagatgaggtgaggtgaggaggaggaagaggaa 1317
 Db 61 CTATAGATGATATGCGGAACACAGAGAGAGACACACTCGATGCTGAGCAGCATGA 4

RESULT 5

ID AAT39517
 XX AAT39517 standard; cDNA; 1209 BP.

AC AAT39517;

XX 12-DEC-1996 (first entry)

XX Flea calreticulin cDNA clone nctCall1209.

XX Calreticulin; flea; haematophagous insect; allergic dermatitis;

KW vaccine; therapy; nctCall1209; ss.

XX Ctenocephalides felis.

XX WO9628469-A1.

XX 19-SEP-1996.

XX 08-MAR-1996; 96WO-US03133.

XX 09-MAR-1995; 95US-0401509.

XX (HESK-) HESKA CORP.

XX Rushlow KE, Stiegler GL;

XX WPI; 1996-442861/44.

DR P-PSDB; AAW04171.

XX

PT Haematophagous insect calreticulin protein - used to reduce insect
 PT infestation and desensitise patients to allergic dermatitis
 XX
 PS Claim 7; Page 71; 86pp; English.
 XX
 CC Nucleic acid nctCall1209 (AAT39517) comprises the coding region (minus
 CC the stop codon) encoding flea calreticulin Pctcal403 (AAW04171), a
 CC protein that can be used to alter the blood feeding behaviour of
 CC haematophagous insects or to desensitise an animal to allergic
 CC dermatitis. The coding sequence was deduced from a cDNA clone
 CC (AAT39516) isolated from a cDNA library prep. from the salivary
 CC glands of Ctenocephalides felis. Calreticulin nucleic acids (see
 CC also AAT39515 and AAT39518-20) can be used to produce recombinant
 CC calreticulins, or can be administered to an animal to protect it
 CC against haematophagous insect infestation.
 XX
 SQ Sequence 1209 BP; 425 A; 212 C; 280 G; 292 T; 0 other;

Query Match 24.6%; Score 481.2; DB 17; Length 1209;
 Best Local Similarity 66.0%; Pred. No. 6.3e-101;
 Matches 712; Conservative 0; Mismatches 363; Indels 3; Gaps 1;

QY 240 agattttggcaaatgtctcagttccggcgaagtctctacgtgacgaggaagaataa 299
 Db 132 agaattcggcaaatgtgcacactgccggaaagtctctataacgagtcggaagcagaa 191
 QY 300 agttttgcagacaagccagatgcacgtctttctctctctctctctctctctctctct 359
 Db 192 agttttgcagacaagccagatgcacgtctctctctctctctctctctctctctct 251
 QY 360 cagcaacaagccagacgctggtgcagttccagttccagttccagttccagttccagtt 419
 Db 252 ctcaataaagaacagacattagttgacaaattttccgttaaacatgaacaaacattga 311
 QY 420 ctgtgtggggcgtgtgtgaaagctgtttcctaatgtttggaccagacagatgcacgg 479
 Db 312 ctgtgtggggcgtgtgtgaaagctgtttcctaatgtttggaccagacagatgcacgg 371
 QY 480 agactcagaatacaacatcatgtttgtgtccagacatctgtgtccctggcaccagaaggt 539
 Db 372 ggaagtcctcatgaatttatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 431
 QY 540 tcattgtcttcaactcaagggcaagacgtgctgatcaacaaaggacacatcctgtgcaa 599
 Db 432 tcactgaattcctcagtcacaagggtaaaaattgtttgatcaataaaggatatacctgtgcaa 491
 QY 600 ggatgatgatttacacacctgtacacactgattgtggcggcagacacacacattatgaggt 659
 Db 492 agatgatgtctatactcatatacactttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 551
 QY 660 gaagattgacaacagccaggtgagtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 719
 Db 552 gttgattgataaaggaggggtgaaagtgttaactgtgaagatgactgtgtgtgtgtgtgt 611
 QY 720 acccaagaagataaaggatcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 779
 Db 612 ccccaagaataaaggatcctgt 671
 QY 780 gatcgatgatccacagactccaagcctgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 839
 Db 672 tattctgacctaatgac 731
 QY 840 ccttgatgtaagaagcccgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 899
 Db 732 tcttgatgtaagaagcccgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 791
 QY 900 agtgattcagaacctgagttacaaaggtgagttggaagcccggtgaggtgaggtgaggtgag 959
 Db 792 tatgattgacaacctgagttacaaaggtgagttggaagcccggtgaggtgaggtgaggtgag 851
 QY 960 ttacaagggcactgtgattccacccagaaattgacacccccgagttatctccgattccacag 1019

Db 852 ctacaaagggtgtctgggttcacccctgaaattgataatccagaatagttctctgatactca 911
QY 1020 tatctatgcctatgataaactttggcgtgctggcctggacctctggcaggtcaagtctgg 1079
Db 912 actttacaaacgtgagagatttggccattggttagattttggaagtaaggctgg 971
QY 1080 caccatctttgacaaacttctctacccaacgtagggcgtacgtagcgtgaggtttggcaa. 1139
Db 972 aacaattatgcacaatttttaataccagatgatgttgattatgcataaagaaatag---c 1028
QY 1140 cagagcgtggcggtacaaaggcagcagagaacaaatgaaggacaaacagcagcagga 1199
Db 1029 agaagtggttaaatctaccaggaaagagaagaaagaaatgaagatgctcaagatga 1088
QY 1200 gcagaggtcttaaggaggaagaaagacaaagacaaagcgaagagggagggagcagagga 1259
Db 1089 agaaaggaaagccagggaagccgaacaaagaaagaaatgacacagatgctgatga 1148
QY 1260 caaggaggatgtaggagacaaagatgaggaaggaggatgagggagacaaagagga 1317
Db 1149 cttagatgataatgccgaacaccagaaagacacactcgatgctgagcagatgaa 1206

RESULT 6
AAT39516
ID AAT39516 standard; cDNA; 1589 BP.
XX
AC AAT39516;
DT 12-DEC-1996 (first entry)
DE
DE Flea calreticulin cDNA clone nCtCall589.
XX
XX Calreticulin; flea; haematophagous insect; allergic dermatitis;
KW vaccine; therapy; nCtCall589; ss.
XX
OS Ctenocephalides felis.
XX
FH Key Location/Qualifiers
FT CDS 151..1362
FT /*tag= a
XX
XX W09628469-AL.
XX
PD 19-SEP-1996.
XX
PF 08-MAR-1996; 96WO-US03133.
XX
PR 09-MAR-1995; 95US-0401509.
XX
XX (HESK-) HESKA CORP.
XX
PI Rushlow KE, Stiegler GL;
XX
XX WPI; 1996-442861/44.
DR P-PSDB; AAW04171.
XX
XX Haematophagous insect calreticulin protein - used to reduce insect
PT infestation and desensitise patients to allergic dermatitis
XX
XX Claim 7; Page 66-68; 86pp; English.
XX
XX A cDNA clone (AAT39516), designated nCtCall589, codes for a flea
CC calreticulin protein (AAW04171) that can be used to alter the blood
CC feeding behaviour of haematophagous insects or to desensitise an
CC animal to allergic dermatitis. Its sequence was deduced from
CC clones nCtCall665, nCtCall750 and nCtCall1218 isolated from a
CC Ctenocephalides felis salivary gland cDNA library by PCR
CC amplification using flea calreticulin gene-specific primers (see
CC also AAT39529-32). nCtCall589 and other calreticulin nucleic acids
CC (see also AAT39515 and AAT39517-20) can be used to produce recombinant
CC calreticulins, or can be administered to an animal to protect it
CC against haematophagous insect infestation.

XX
SQ Sequence 1589 BP; 552 A; 266 C; 335 G; 436 T; 0 other;
Query Match 24.6%; Score 481.2; DB 17; Length 1589;
Best Local Similarity 66.0%; Pred. No. 7.1e-101;
Matches 712; Conservative 0; Mismatches 363; Indels 3; Gaps 1;
QY 240 agattttggcaaatcttctcagttccggcaagttctcagtgacgaggaagataa 299
Db 282 agaatcggcaaatctcgtgcacactccggaagttctataacgataccgaagcagaaa 341
QY 300 agtttgcagacaagccagggatgcacgctttttatgtctctgcggccagtttcagccttt 359
Db 342 agtttgcacaacaggtcgaagtgtctcagctctatcgtctcatttaagttcaaaccttt 401
QY 360 cagcaacaaggccagcagcgtggtgcagttcacggtgaaacatgacagacaatcga 419
Db 402 ctcaaataaagacagacattggtgacaatttcttcgttaaacatgaacaaacattga 461
QY 420 ctgtggggcgcgctatgtgaagctgtttccttaattgttgaccagacagatgcacgg 479
Db 462 ctgtggaggtggttacttgaaaggtttcgaattcagtgatgaatcaaaagacatgcattg 521
QY 480 agactcagaatacaacatcatgtttgtcccgcacatctgtggccctggccaagaaggt 539
Db 522 ggaagtcctctatgaaatttatgttgcctgacatttgcaccaggaaactaagaaggt 581
QY 540 tcatgtcatcttcaactacaagggaagacgctgctgatcaacaaggacatccgtttcaa 599
Db 582 tcagtaattctcagtcacaggtgtaaaatgtttgatcaataaggtatccctcgctcaa 641
QY 600 ggatgaggtttacacacctgtacacactgattgtgcccagacacaacacctatgaggt 659
Db 642 agatgagtctatactcatgtatacactttggtgttaagcccgataatacctatgaggt 701
QY 660 gaagattgacaacagccaggtggagtcctcctctctggaagacgattgggacttccctgcc 719
Db 702 gtgattgataatgagaaggttgaaagtgttaacttggaaagatgactgggaattcctagc 761
QY 720 acccaagaataaaggatcctgattcctcaaaaccggaagactggatgagcgggcca 779
Db 762 ccccaagaaatcaaggtatccagaagctaaataaaaccagcagattgggtatgagcgtgcaac 821
QY 780 gatcgatgccacacagctccagcctgaggtgggactgggacagcccgagatataccctga 839
Db 822 tatctgacctaatgacacccaaacctgaagattgggacaaagccagaacacattcctga 881
QY 840 ccctgatgctaagaagcccgagactgggtgaagagatggacggagtgaggagccccc 899
Db 882 tectgatctaccaaacctgatgactgggtgaagagatggatggtgaaagggaacctgc 941
QY 900 agtgattcagaacctgagtacaaaggtgagtggaagcccccggcagatcgacacaccaga 959
Db 942 tatgattgacaacctgaatataaaggagaaatgggcccacaaacagattgacatacctga 1001
QY 960 ttacaaggcacttggtatccaccagaaattgacaccccgagtattctcccgatccag 1019
Db 1002 ctcaaaagggtgtctgggttccacctgaaattgataatccagaatatgttctctgatactca 1061
QY 1020 tatctatgcctatgataaactttggcgtgctggcctggacctctggcaggttcaagtctgg 1079
Db 1062 actttacaaacgtgagatttggccattggtttagattttggaagtaaggctgg 1121
QY 1080 caccatcttgacaacttctctacccaacgtaggagcgtacacgtaggaggtttggcaa 1139
Db 1122 acaaatctgcacaataattttaatcacagatgattgttattatgataaagaaaaatag---c 1178
QY 1140 cgagacgtgggggtataaaggcagcagagaacaaatgaaggacaaacagcagcagga 1199
Db 1179 agaaggtgttaaatctcaccaggaaaggaagaaagaaatgaagatgtcgaagatga 1238
QY 1200 gcagaggtcttaaggaggtggaagaaagacaaagacgaagaggtgaggaggtgagagga 1259

```

Db 1239 agaaaggaaagccagggaagcgaacaaagaaagaaatgacacagatgctgatgaaga 1298
Qy 1260 caaggaggatgagagacaaagatgagatgaggagatgaggagagacaaaggagaa 1317
Db 1299 cttgatgataatgccgaacaccagaaagaaagacacactcgatgctgagcacagatgaa 1356

RESULT 7
AAT39518/c
ID AAT39518 standard; cDNA; 1589 BP.
XX
AC AAT39518;
XX
DT 12-DEC-1996 (first entry)
XX
DE Flea calreticulin cDNA clone nctCall1589 (complement).
XX
KW Calreticulin; flea; haematophagous insect; allergic dermatitis;
KW vaccine; therapy; nctCall1589; ss.
XX
OS Ctenocephalides felis.
XX
PN WO9628469-A1.
XX
PD 19-SEP-1996.
XX
PF 08-MAR-1996; 96WO-US031133.
XX
PR 09-MAR-1995; 95US-0401509.
XX
PA (HESK-) HESKA CORP.
XX
PI Rushlow KE, Stiegler GL;
XX
WPI; 1996-442861/44.
XX
DR
XX
PT Haematophagous insect calreticulin protein - used to reduce insect
PT infestation and desensitise patients to allergic dermatitis
XX
PS Claim 16; Page 70; 86pp; English.
XX
CC The antisense strand (AAT39518) of nucleic acid nctCall1589 (AAT39516)
CC hybridises under stringent conditions to a flea calreticulin gene.
CC nctCall1589 codes for a flea calreticulin protein (AAW04171) that can
CC be used to alter the blood feeding behaviour of haematophagous
CC insects or to desensitise an animal to allergic dermatitis. The
CC antisense strand, when administered to an animal, reduces
CC calreticulin activity in insects feeding on the animal, thereby
CC reducing the insect burden on the animal and in the environment
CC of the animal.
XX
SQ Sequence 1589 BP; 436 A; 335 C; 266 G; 552 T; 0 other;

Query Match 24.6%; Score 481.2; DB 17; Length 1589;
Best Local Similarity 66.0%; Pred. No. 7.1e-101;
Matches 712; Conservative 0; Mismatches 363; Indels 3; Gaps 1;

Qy 240 agatttggcaaatcgttctcagttccggaagttctacggtgacgaggagaagataa 299
Db 1308 AGAATTGCGCAAAATCGTGCACACTGCCGGAAGTTCTATAACGATGCCGAGCAGACAA 1249
Qy 300 aggtttgcagacaaagccagagtcacgcgttttatgctctgctcgccagtttcgaccttt 359
Db 1248 AGGTTTGCACAAAGCTCAAGATGCTAGTGTCTACGCTCTATCTCTATAAGTTCAACCTTT 1189
Qy 360 cagcaacaaagccagacgcgtggtgtgagttcaggttcaacgtgaaacatgagcgaacatoga 419
Db 1188 CTCAAATAAAGACACAGACATAGTTGTACAATTTTCCGTTAAACATGAACAAACATTGA 1129
Qy 420 ctgtggggcggtctatgtgaagctgtttccctaataagtttgaccagacacatgcacgg 479

```

```

Db 1128 CTGTGGAGGTGTTACTTGAAGGGTTTCGAATTCAGTGTGAATCAAAAGGACATGTCATGG 1069
Qy 480 agactcagaatacaacatcatgtttgtcccgacatctgtgcccctggtgcaaccaagaaggt 539
Db 1068 GGAAGTCCCTATGAATATATGTTTGTCTGCTGACATTTGTGACCCAGGAACCTAAGAAGGT 1009
Qy 540 tcatgtcatcttcaactacaaggcgaagaacgtgtcgatcaacaagacacccgttgcacaa 599
Db 1008 TCACGTAATCTTCAGCTACAAAGGGTAAATGTTTGTATCAATAAGGATATCCGCTGCAA 949
Qy 600 ggatgatgagtttacacactgttacacactgtgtgctggtgacacacacactatgaggt 659
Db 948 AGATGATGCTTACTACTCATGTATACACTTTGTGTTGTTAAAGCCCGATAATACCTATGAGGT 889
Qy 660 gaagattgacaacacccaggtgagtcctggtcctcttgaagacgattggagacttctgtcc 719
Db 888 GTTGATTGATATGAGAAGGTGAAAGTGGTAACCTTGGGAAGTACTGGGAATTCCTAGC 829
Qy 720 acccaagaagataaaggatcctgatgcttcaaaacccggaagactggtgagtcgagcgccaa 779
Db 828 CCCCAAGAAATCAAGGATCCAGAGCTAAATAAACCCAGCAGATGGGATGAGCGTCAAC 769
Qy 780 gatcgatgattccacagactccaaagcctgagactggtggaagcccgagatccctga 839
Db 768 TATTCCTGACCTTAATGACACCAAACTGAAGATTGGGACAAAGCCAGAACACATTCCTGA 709
Qy 840 ccctgatgctaagaagcccgagactggtggtgaagatgagcggagagtggaagaccccc 899
Db 708 TCCTGATGCTACCAACCTGATGACTGGGTGAAGAGATGGATGGTGAATGGGAACCTGC 649
Qy 900 agtattcagaacccctgagtacaaaggtgagtggaagcccccgcagatcgacacaccaga 959
Db 648 TATGATTGACAACCCCTGAATATAAGGGAGAATGGGCCACCAAAACAGATTGACAATCCTGA 589
Qy 960 ttacaagggcacttgatccaccagaaattgacaccccgagtagttctcccgatccccag 1019
Db 588 CTACAAAGGTGCTGGGTTACCTGAAATTCAGATATCCAGAATATGTTCTGTACTCA 529
Qy 1020 tatctatgctatgataactttggcgtgctggcctggaacctctggcaggtcgaagctcgtg 1079
Db 528 ACTTTACAAACGTGATGAGATTGTGCCATTGTTAGATTATGGCAAGTAAAGGCTGG 469
Qy 1080 caccattttgacaacttctcatcaccaacgatgagcgatcgctgagagagtttggcaa 1139
Db 468 AACAAATATTCGACAATATTTTAAATCACAGATGATGTTGATTATGCAGAAATAG---C 412
Qy 1140 cgagacgtggtggcggtacaaagggcgagagaaacaaatgaaggacaaacagagcagga 1199
Db 411 AGAAGGTGTTAAATCTACCCAGGAAGAGAAAGAAATGAAAGATGCTCAAGATGAAGA 352
Qy 1200 gcagaggttaagggaggaagaagacaagaacgcaagagaggagagagagcagagga 1259
Db 351 AGAAGGAAGCCGCGGGAAGCGGAAACAAAGAAAGAAATGACAGATGCTGTATGAAGA 292
Qy 1260 caaggaggatgatgaggacaaagatgaggatgaggagatgaggagaggaagagaa 1317
Db 291 CTTAGATGATAATCCGAAACACACAGAGAGACACACTCGATGCTGAGCAGATGAA 234

RESULT 8
AAQ11987
ID AAQ11987 standard; DNA; 1076 BP.
XX
AC AAQ11987;
XX
DT 29-AUG-1991 (first entry)
XX
DE Insert from lambda RAL-1 encoding 42 kD antigen.
XX
KW River blindness; onchocerciasis; vaccine; antigen; parasite; ss.
XX
OS Onchocerca volvulus.
XX

```

FH Key Location/Qualifiers
 FT CDS 1..1008
 FT /tag= a
 FT /label= partial sequence of O. volvulus antigen
 FT /note= "5' terminal absent"
 FT repeat_unit 478..498
 FT /tag= b
 FT repeat_unit 529..549
 FT /tag= c
 FT repeat_unit 580..600
 FT /tag= d
 FT polyA_signal 1035..1040
 FT /tag= e
 FT polyA_signal 1044..1049
 FT /tag= e
 XX US021342-A.
 XX
 PD 04-JUN-1991.
 XX
 PF 30-JUN-1988; 88US-0214264.
 XX
 PR 30-JUN-1988; 88US-0214264.
 XX
 PA (UYHO-) UNIV HOSPITALS CLEV.
 XX
 PI Greene BM, Unnasch TR;
 XX
 DR WPI; 1991-185179/25.
 DR P-PSDB; AARI2312.
 XX
 PT DNA encoding Onchocerca volvulus antigen - used to express
 PT recombinant antigen for vaccine against onchocerciasis or river
 PT blindness.
 XX
 PS Claim 1; Fig 7; 20pp; English.
 CC
 CC The sequence was obtd. from cDNA clone lambda RAL-1 prep'd. from
 CC RNA isolated from nodules excised from patients infected with O.
 CC volvulus. A library prep'd. from the RNA was screened with anti-
 CC sera prep'd. from O. volvulus infective larvae. The insert has a
 CC single open reading frame but does not include the entire 5' end.
 CC However, Abs which specifically bind to protein prep'd. from in-
 CC duced cultures of lambda RAL-1 lysogens recognise a single poly-
 CC peptide of mol. wt. 42,000 in extracts of adult worms. Analysis
 CC of the ORF suggests that it encodes a protein of mol. wt. 39,130.
 CC If the antigen is not subject to post-translational processing
 CC this suggests that most of the coding sequence is present. The
 CC three repeats encode highly hydrophilic regions likely to be ex-
 CC posed on the surface of the antigen and highly immunogenic.
 CC Recombinant antigen expressed by the clone can be used stimulate
 CC T-cells of individuals infected by the parasite to proliferate
 CC and may be used as the basis for a vaccine against Onchocercia-
 CC sis or river blindness.
 XX
 SQ Sequence 1076 BP; 435 A; 153 C; 258 G; 232 T; 0 other;

Query Match 22.4%; Score 438.8; DB 12; Length 1076;
 Best Local Similarity 66.3%; Pred. No. 3.2e-91;
 Matches 662; Conservative 0; Mismatches 332; Indels 4; Gaps 2;
 QY 274 ttctacgtgacgaggagaagaataaagggttgcagacaagccagagctggtgtgcag 333
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1 ttttatgcatgctgtgaagataaagggtctcaagacaacacagagtgcaaatattat 60
 QY 334 gctctgtcggccagtttgcg---agcctttcagacaacaaagccagagctggtgtgcag 390
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 61 tcgatcgtgcaaatgtgataagattgttcaaacagggtaaatcgtggttatccaa 120
 QY 391 ttacacgtgaaacatgacgaacatgactgtggtggcggtctatgtgaagctgttctc 450
 Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 QY 121 tttagctcaagcatgacaggagagattgattgtggtggtggtggttaagctaatggca 180

QY 451 aatagtttgaccagacagacatgacgagagactcagaatacaaacatcatgtttggtccc 510
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 181 tcggtatgaacctagaaagtattcccatggtgaaactccttatcatatcatgttcggtcc 240
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 511 gacatctgtgcccctggcaccagaagaagtttcatgtctatcttcaactacaaggggcaagaac 570
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 241 gacatctgtgacctggacaagaagaagtcatttccattataaaggacagagaat 300
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 571 gtgctgatcaacaaggacatccgttgaaggtatgatttacaacacctgtcacactg 630
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 301 catatgatcaagaagacattcgtgcaaggatgattcttcaactctctcacacttg 360
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 631 attgtcgccgacacacacctatgaggtgaagattgacaacagcccggtgagtcgggc 690
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 361 attgtaacctgacaacacctatgaatttcaaatgacggtgaaaaggtgaatcctggt 420
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 691 tccctggaagacgattgggacttccctcccccagaagaataaaggatccttgatgcttca 750
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 421 gagctggaagcagattgggacttctctccaaagaaaattaaaggatccagatgcaaaa 480
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 751 aaaccggaagactgggatgagcgggccaagatcgatccacagactccaagcctgag 810
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 481 aaaccgaagactgggacgagcgaattttatcgtgatgaggtatgaagaaccggaa 540
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 811 gactgggacaagcccgagcatatccctgacctgatgctaagaagcccgagactgggat 870
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 541 gattgggacaagccggaacatttcgtgacctgatgctaagaacccggagactgggat 600
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 871 gaagagatgacggagagtggtggaaaccccgatgattcagaacccctgagtacaaggggtgag 930
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 601 gatgaatggtggtggatgggagcccgatggttgataatcctgaataaaaggagag 660
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 931 tggaaagcccgcgagatcgacaacccagattacaaggcgacttggtccacccagaatt 990
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 661 tggaaagccgaacaaagaaaaatcctgcatacaaggcgaattggatccatccagaaatc 720
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 991 gacaaccccgagtattctccgactcccgatctatcctcatgataactttggcggtg 1050
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 721 gaaattccagattatactccagacgataaattgtgtattgttatgatattgagacatt 780
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1051 ggcttgacctctgacgaggtcaagtctggaccaccttttgacaacttctcatcaccaac 1110
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 781 ggcttgatttggcgaagtaaatcaggaacctttttgtatgcgtcatgtgacagat 840
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1111 gatgagcatagctgaggagtttggaacagagacgtggggcgtaacaaaaggcagcagag 1170
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 841 agtgtgaggaggccaagaagtttggtaaaagacattgaaataacagaggggaaggtgaa 900
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1171 aaacaaatgaaggacaacacagagcagagcagaggttaaggaggaagaagacaag 1230
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 901 aagaaaagggaacaaaacaaaacagagaagaaaaggaggaatgaa-aagatcaa 959
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1231 aaacgcaaaaggaggaggagcagagacagagga 1268
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 960 aaaggaaaaaatgaaaaaaaggaaaaagacagaa 997
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
 AAC02428
 ID AAC02428 standard; cDNA; 385 BP.
 XX
 AC AAC02428;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 2426.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX

PN EP1033401-A2.
 XX 06-SEP-2000.
 XX 21-FEB-2000; 2000EP-0200610.
 PF 26-FEB-1999; 99US-0122487.
 PR (GEST) GENSET.
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 DR P-PSDB; AAG02422.
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 2426; 71pp + CD-ROM; English.
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.
 XX Sequence 385 BP; 115 A; 76 C; 124 G; 56 T; 14 other;
 SQ
 Query Match 18.9%; Score 369.8; DB 21; Length 385;
 Best Local Similarity 95.1%; Pred. No. 1.4e-75;
 Matches 365; Conservative 13; Mismatches 6; Indels 0; Gaps 0;
 QY 963 caagggcacttgatccaccggaattgacaccccgagttatctcccgatcccgat 1022
 DB 1 caaggcacttgatccaccggaattgacaccccgagttatctcccgatcccgat 60
 QY 1023 ctatgcctatgataactttggcgtgctggcctgacacctgacgtcgaagtctggcac 1082
 DB 61 ctatgcctatgataactttggcgtgctggcctgacacctgacgtcgaagtctggcac 120
 QY 1083 catcttgacaacttctctcatcaccacagatgagcctacgtcgtgaggtttggcaacga 1142
 DB 121 catcttgacaacttctctcatcaccacagatgagcctacgtcgtgaggtttggcaacga 180
 QY 1143 gacgtggggtgtaacaaaggcgacgagagaacaaatgaaggacaaacagcagagagca 1202
 DB 181 gacgtggggtgtaacaaaggcgacgagagaacaaatgaaggacaaacagcagagagca 240
 QY 1203 gaggcttaaggagaggaagacacaaacgcaaaagagagagagagagagagagca 1262
 DB 241 gaggcttaaggagaggaagacacaaacgcaaaagagagagagagagagagagagca 300
 QY 1263 ggaggatgatgaggacaaagatgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 1322
 DB 301 ggaggwtgatgagmcaagwtgagwtgagwtgagwtgagwtgagwtgagwtgagwtgag 360
 QY 1323 ggaggaaatgtctcccgccaggcc 1346
 DB 361 ggnagaagatgtcccgcmaggcc 384
 RESULT 10

AAX27251
 ID AAX27251 standard; DNA; 366 BP.
 XX
 AC AAX27251;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE Human cClqR binding domain coding sequence.
 XX
 KW Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
 KW CUB functionality; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9907406-A1.
 XX
 PD 18-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-GB02430.
 XX
 PR 12-AUG-1997; 97GB-0016998.
 XX
 PA (UYLE-) UNIV LEICESTER.
 XX
 PI Schwaebler W;
 DR WPI; 1999-180404/15.
 DR P-PSDB; AAY00924.
 XX
 PT Use of a cClqR binding domain - to modulate complement ubiquitin
 PT (CUB) functionality.
 XX
 PS Disclosure: Page 25; 31pp; English.
 CC This sequence encodes a Clq and collectin receptor (cClqR) binding
 CC domain. The invention relates to the use of a cClqR binding domain in a
 CC medicament to effect complement ubiquitin (CUB) functionality, and an
 CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 CC functionality. The cClqR binding domain, or its inhibitor, can be used to
 CC treat a human or animal body. Particularly an inhibitor is used to treat
 CC complement activation involved in the initiation and maintenance of
 CC inflammation, for example in myocardial infarction, brain ischaemia
 CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 CC erythematosus, burns, immune complex nephritis, and to treat amyloid
 CC plaques in Alzheimer's disease. The use of cClqR binding domain or
 CC inhibitor enables the CUB domain functionality to be modulated using a
 CC low molecular weight molecule.
 XX
 SQ Sequence 366 BP; 107 A; 92 C; 107 G; 60 T; 0 other;
 Query Match 18.7%; Score 366; DB 20; Length 366;
 Best Local Similarity 100.0%; Pred. No. 9.9e-75;
 Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 592 cgttccaaggatgatgatttaccacacctgtacacactgattgtcgcgcacacacacc 651
 DB 1 cgttccaaggatgatgatttaccacacctgtacacactgattgtcgcgcacacacacc 60
 QY 652 tatgaggtgaagatgacaacgccaggtggagtcgctcttgaagacattggac 711
 DB 61 tatgaggtgaagatgacaacgccaggtggagtcgctcttgaagacattggac 120
 QY 712 ttctgtccaccacaaagaataaaagatcctgatcttcaaacccggaagactggatgag 771
 DB 121 ttctgtccaccacaaagaataaaagatcctgatcttcaaacccggaagactggatgag 180
 QY 772 cgggccaaagatcgatgccacacagactccaaagctcgaagctgggacagcccgagcat 831
 DB 181 cgggccaaagatcgatgccacacagactccaaagctcgaagctgggacagcccgagcat 240

RESULT 13
 AAF29755
 ID AAF29755 standard; cDNA; 1513 BP.
 XX
 AC AAF29755;
 XX
 DT 05-APR-2001 (first entry)
 XX
 XX Castor bean calreticulin coding sequence.
 DE
 DE Castor bean; calreticulin; calcium binding; er; endoplasmic reticulum;
 KW calnexin; promoter; resistance; ss.
 KW
 XX
 OS Ricinus communis.
 XX
 XX US6171864-B1.
 PN
 XX 09-JAN-2001.
 PD
 XX 05-JUL-1996; 96US-0675816.
 PF
 XX 05-JUL-1996; 96US-0675816.
 PR
 XX (PION-) PIONEER HI-BRED INT INC.
 PA
 XX Coughlan SJ, Winfrey RJ;
 PI
 XX WPI: 2001-122335/13.
 DR
 DR P-PSDB; AAB66343.
 XX
 PT New nucleic acid molecules encoding a calcium binding chaperone protein
 PT in endoplasmic reticulum, calreticulin, and calreticulin promoter
 PT sequences, useful for producing foreign gene products in plant cells -
 XX
 PS Disclosure; Fig 1A; 45pp; English.
 XX
 CC The present invention provides the protein and coding sequences for the
 CC castor bean calreticulin protein and the calreticulin promoter sequence.
 CC In addition, the castor bean calnexin protein, coding sequence and
 CC promoter are also described. Calreticulin and calnexin are calcium
 CC binding proteins found in the endoplasmic reticulum (er). The
 CC calreticulin promoter sequence can be used in vectors to promote the
 CC expression of foreign genes, particularly resistance genes, in plant
 CC cells.
 XX
 SQ Sequence 1513 BP; 504 A; 268 C; 363 G; 378 T; 0 other;

Query Match 18.4%; Score 360; DB 22; Length 1513;
 Best Local Similarity 60.4%; Pred. No. 4.2e-73;
 Matches 635; Conservative 0; Mismatches 405; Indels 12; Gaps 2;

Qy 282 tgacgagagagaataaaggttgacagacagccagcagcagcgttttgcctgtc 341
 Db //
 183 tggagaccctaatgacaaaggtattcaacaaagtgaaagattataggtttctatcttc 242
 Qy 342 ggcagtttcgagcctttcagacacaaaggccagacgctggtgtgagttcagcgggtgaa 401
 Db //
 243 agctgaattcctgaattcagataaagataagactctagctcttccaaattttctgcaa 302
 Qy 402 acatgagacagacacgactgctggtggcggtctatgtgaagctgtttctccaatagtttga 461
 Db //
 303 gcatgaacagaagttgactgt 362
 Qy 462 ccagacagacacgacgagactcagatacaacatcatgttttgggtccgacatctgtgg 521
 Db //
 363 ccagaagaatttgggtgtgacactccacagatcatgttttgggtgtgtgtgtgtgtgtgtgtgt 422
 Qy 522 ccctggcaccagaaggtttcatcttcattcattcattcattcattcattcattcattcattc 581
 Db //
 423 ctacagcaccacaaaggttcatgtctatcttcaactacaatgatacaaacaccattgatcaa 482

Qy 582 caagagacatccgttgcaaggatgatgagtttacacacacgttacacactgattgtcgcc 641
 Db //
 483 aaaggaaagtccatgtgaaacccgacagcttaactatgttttacacattggtcatccgcc 542
 Qy 642 agacaacacattatgaggtgaagattgacaacagccaggtgagtcggtcctcttggaaga 701
 Db //
 543 agatgctacttatagcattcttcgacaatgtggaagacaaactgggtatgtgtacac 602
 Qy 702 cgattggacttctccaccacaaagaataaagatcattcattcattcattcattcattcattc 761
 Db //
 603 tgaatgggattcttctccaccctaaagaaattaaaggaccttgagccaaagaacagaga 662
 Qy 762 ctgggatgagcggcccaagatcgatgccacagactcccaagcctgagcgtgagcactgg--ga 818
 Db //
 663 ttgggatgagaagagtatatctcctgacctgaggataaagaaccagaggtgtatgatga 722
 Qy 819 caagcccgagcatatccctgacctgacctgacctgacctgacctgacctgacctgacctgacct 878
 Db //
 723 cattccaaaggagattccagatcccgatgccagaaagcctgagtcacaaaggtggtgatgagga 782
 Qy 879 ggacggagagtggaaaccccgatgattcagaacccctgagtcacaaaggtggtgaggaagcc 938
 Db //
 783 agatggtgaaatggactgcccaaccattgccaacccctgagtcacaaaggtggtgatgagga 842
 Qy 939 ccggcagatcgacacacccagattacaaggcacttggatccaccacagaaattgacaaccc 998
 Db //
 843 caagaaattaaagaaccccaactacaaggcaagtggaaagcaccatgatcgacaaccc 902
 Qy 999 cgatttctccgacatccacagtatctatctatctatctatctatctatctatctatctatctatct 1058
 Db //
 903 agattcaaggatgaccacagaaattctgttaccacactgaagtgtgtgtgtgtgtgtgtgtgt 962
 Qy 1059 cctgtcagatcaagtctgcacacattcttgacaaactctctcatccacacagatgagggc 1118
 Db //
 963 attggtgaggtgaaatctggaactctgttgcaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1022
 Qy 1119 atagcgtgagaggttggaacagagcgtggtggcggtgaacaaaggcagagagaacaaat 1178
 Db //
 1023 gtatgtaagcagctggtgagagacatgggaaagacaaagatgctga----- 1073
 Qy 1179 gaaggacaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1238
 Db //
 1074 gaagc 1133
 Qy 1239 agagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1298
 Db //
 1134 agctgattctgtgctgacggagcagtgatgagtgatgagtgatgagtgatgagtgatgagtgatgag 1193
 Qy 1299 tgaggagacaaaggag 1330
 Db //
 1194 tggtaagcaaatcagtgacagcagagacag 1225
 RESULT 14
 AAC40452
 ID AAC40452 standard; DNA; 1532 BP.
 XX
 AC AAC40452;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 28348.
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 XX 06-SEP-2000.
 XX

PF 25-FEB-2000; 2000EP-0301439.
XX 99US-0121825.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149457.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 16:48:06 ; Search time 59.69 Seconds
(without alignments)
6209.938 Million cell updates/sec

Title: US-09-807-148-7
Perfect score: 1958
Sequence: 1 gaattccggtatctttaa.....aaaaaaaaaaaggatttc 1958

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	363.4	18.6	1514	4	US-08-675-816-1
2	108.6	5.5	3191	4	US-08-675-816-3
3	97.8	5.0	3489	2	US-08-728-323A-1
C 4	97.8	5.0	32207	2	US-08-770-379-20
C 5	97.8	5.0	32207	4	US-08-757-669A-20
C 6	97.2	5.0	1931	2	US-09-130-114-2
C 7	96.6	4.9	7218	1	US-08-232-463-14
C 8	95.4	4.9	51259	3	US-08-781-891-209
C 9	91	4.6	16442	3	US-08-781-891-208
10	84.2	4.3	7218	1	US-08-232-463-14
11	83.4	4.3	3100	1	US-08-296-362-1
12	79.6	4.1	3211	2	US-08-574-959A-8
13	79.6	4.1	3901	2	US-08-574-959A-6
14	78.8	4.0	2301	1	US-08-306-691B-23
15	78.8	4.0	2301	5	PCT-US93-06251-78
16	77.4	4.0	2580	3	US-09-050-863-2
C 17	77.4	4.0	5452	2	US-09-130-114-1
18	77.4	4.0	9600	4	US-08-910-647-1
19	77.4	4.0	10596	1	US-07-884-811-15
20	77.4	4.0	10596	1	US-07-885-971-15
21	77.4	4.0	10596	1	US-08-087-783A-15
22	77.4	4.0	10596	1	US-08-194-088B-15
23	77.4	4.0	10596	2	US-08-194-087-15
24	77.4	4.0	10596	5	PCT-US93-04648-15
25	74.8	3.8	390	4	US-09-197-649-7
26	73	3.7	2072	4	US-08-675-816-5
27	73	3.7	4358	4	US-08-675-816-7
28	70.6	3.6	2518	4	US-09-433-699-3
29	68.4	3.5	120	2	US-08-114-555A-4
30	68.4	3.5	120	3	US-08-559-397A-6
C 31	68	3.5	2188	1	US-07-865-662F-10
C 32	68	3.5	2188	4	US-08-374-219B-10
33	67.2	3.4	2277	1	US-08-676-967-2
34	67.2	3.4	2277	1	US-08-676-974-2
35	67.2	3.4	2277	2	US-09-098-487-2
36	66.8	3.4	120	3	US-08-114-555A-5
37	66.8	3.4	120	3	US-08-559-397A-8
38	66.8	3.4	43795	3	US-08-742-185-101
39	65	3.3	2295	1	US-08-375-300-3
40	65	3.3	2295	3	US-09-177-431-3
41	65	3.3	2295	5	PCT-US95-16930-3
42	65	3.3	4080	1	US-08-375-300-1
43	65	3.3	4080	3	US-09-177-431-1
44	65	3.3	4080	5	PCT-US95-16930-1
45	63.8	3.3	2340	3	US-09-022-983-4

RESULT 1
US-08-675-816-1
; Sequence 1, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND US
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)-622-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1514 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1254
; US-08-675-816-1

Query Match 18.6%; Score 363.4; DB 4; Length 1514;
Best Local Similarity 60.4%; Pred. No. 1.4e-77;
Matches 658; Conservative 0; Mismatches 416; Indels 15; Gaps 3;

Qy	282	tgacgaggagaagataaaagttgtagacaagccaggatgcacgcttttatgtctctgc	341
Db	183	TGGAGACCTAATGACAAAGGTATTCAACAACAGTGAAGATTATAGTGTCTATGCTATTTC	242
Qy	342	ggccagtttcgagcccttcagacaacaaaggccagacgctgggtgcagttcaagtgtaa	401
Db	243	AGCTGAATTCCTCGAATTCAGTAAATAAAGATAGACTCTAGTCTTCCAAATTTCTGTCAA	302
Qy	402	acatgacgagaacatgactgtggggcgctatgtgaagctgtttcctaatagtttgga	461
Db	303	GCATGAACAGAGCTTGACTGTGGTGGTTACATGAAGTTGCTCAGTAGTAGCACTGA	362
Qy	462	ccagacagacatgcaocggagactcagaatacaacatcatctgtttgggtcccgacatctg	521
Db	363	CCAGAAGAAATTTGGTGGTGACACTCCATACAGTATCATGTTTGGACCCGATATATGTGG	422
Qy	522	ccctggcaccagaagttcattcatcttcactacaaggccaagaacgctgcgtatcaa	581
Db	423	CTACAGCACCAAAAGGTTTCATGCTATCTTAACTACATGAATACAAACCACTTGATGCAA	482
Qy	582	caagacatccgttgaagatgatgagtttacacacctgtacacactgattgicgggcc	641
Db	483	AAAGGAAGTTCCATGTGAACCCGACCAGTTAACTCATGTTTACACATTTGGTCATCCGTCC	542
Qy	642	agacaacacctatgaggtgaagattgacaacagccaggctggagctccggtccctctgga	701
Db	543	AGATGCTACTTATAGCATTTCTTATCGACAAATGTGGAGAGCAAACTGGTGTGTGTACAC	602
Qy	702	cgattggacttcctgccaccaagaagataaagatcctgatcctcaaacccgaaga	761
Db	603	TGACTGGGATCTTCTCCACTTAAGAAAATTAAGGACCTTGAGGCCAGAAACCAAGAAGA	662
Qy	762	ctgggatggcggccaagatgatccccacagactccaagcctgagg---actggga	818
Db	663	TTGGGATGAGAAGGAGTATTCTTGACCTTGAGGATTAAGAAACCAGAGGGTTATGATGA	722
Qy	819	caagccgagcatatccctgacctgatctgaagaagcccgaggactgggtgaagagat	878
Db	723	CATTCCAAAGGAGATTCCAGATCCCGATGCCAAGAAGCCTTGAGGATTTGGGATGATGAGA	782
Qy	879	ggacggagagtggaaaccccgatgattcagaacctgagtacaaaggtgagtgaagcc	938
Db	783	AGATGGTAATGAGACTGCCCAACCAATGCCAACCCCTGAGTACAAGGTTCCATGGAACC	842
Qy	939	ccggcagatcgacaacccagattacaaggcgacttggatccccaccagaattgacaacc	998
Db	843	CAAGAAAATTAAGAACCCCAACTACAAAGGCAAGTGGAAAGCACCACCAATGATCGACAACC	902
Qy	999	cgagtattcccgatcccgatctatgctctatgataaaccttggcgtgctgggccttgg	1058
Db	903	AGATTTCAGGAGTGACCCAGAAATCTATGTTTACCCCAACTTGAAGTATGTTGGTATTGA	962
Qy	1059	ccctggcaggtcaagctctggcaccatcttgacaacttcctcatccaacgatgaggc	1118
Db	963	ATTGTGGCAGGTGAATCTTGGAACTTGTGTTGACAAATGTCTTGATTTCATATGACCCCTGA	1022
Qy	1119	atacgtgaggagtttggcaacgagacgtggggcgtaacaaaggcagcagagaacaaat	1178
Db	1023	GTATGCTAAGCAGCTGGCTGAAGAGACATGGGGAAGAACAAAGATGCTGA-----	1073
Qy	1179	gaaggacaacaggaacgaggcagagagccttaaggaggaggagaagaagaacgaacgcaa	1238
Db	1074	GAAGGCCAGCATTTGAAGAGCGCAGAGAAGAAGAAGAAGGAGGAATCAAGAGTATGATCC	1133
Qy	1239	agaggaggaggcagagacaagagagatgatagacaaaagataggatgaggaga	1298
Db	1134	AGCTGATTCTGATCTGCAGAGGACGATGATGCTGATGNACTGAAGGAGAAGATGA	1193
Qy	1299	tgaggaggacaaggagaagaatgagaggagaagatgtccccggccaggccaaggcagagct	1358
Db	1194	TGGTGAAGAACA---AATCAGATTCGACGACAAGACACAGTCTGAGGACGTACATGATGAAT	1250

Qy 1359 gtagagagg 1367
 |||
 Db 1251 GTAGAGAGG 1259

RESULT 2

US-08-675-816-3
 ; Sequence 3, Application US/08675816
 ; Patent No. 6171864
 ; GENERAL INFORMATION:
 ; APPLICANT: Coughlan, Sean J.
 ; APPLICANT: Winfrey, Jr., Ron J.
 ; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Seed and Berry
 ; STREET: 701 Fifth Ave. Suite 6300
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: U.S.A.
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/675.816
 ; FILING DATE: 05-JUL-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 6171864tenburg, Carol
 ; REGISTRATION NUMBER: 39,317
 ; REFERENCE/DOCKET NUMBER: 750027.401
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206-622-4900
 ; TELEFAX: (206)-682-6031
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3191 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-675-816-3

13

[illegible]

[illegible]

Db 1266 YY 1325
Qy 1757 ggtgtgtctccaaacccagcactgaggaagacgggctctctctcattccaccctc 1816
Db 1326 YY 1385
Qy 1817 cctttccctccgcccccagcagtgggccacttctggtggggcagtggtgtccagatt 1875
Db 1386 YY 1444

RESULT 11
US-08-296-362-1
; Sequence 1, Application US/08296362
; Patent No. 5691306
; GENERAL INFORMATION:
; APPLICANT: Bergeron, John J.M.
; APPLICANT: Thomas, David Y.
; APPLICANT: Wada, Ikuo
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; TITLE OF INVENTION: PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
; TITLE OF INVENTION: PROTEIN PRODUCTION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,362
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deehr, Marya S.
; REGISTRATION NUMBER: 37,120
; REFERENCE/DOCKET NUMBER: 690066.401c1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 102..1883
US-08-296-362-1

Query Match 4.3%; Score 83.4; DB 1; Length 3100;
Best Local Similarity 63.0%; Pred. No. 1e-10;
Matches 192; Conservative 0; Mismatches 86; Indels 27; Gaps 3;
Qy 727 aagataagatcctgctgattcttaaaacccggaagactgg-----gatgagcgggccaag 780
Db 978 AAAATACAGATCTGTGCTGTCATCAACACGATGACTGGAATGAAGATGCCCTGCTTAAG 1037
Qy 781 atcgatgccacagactccaagcctgagactgg-----gacaagcccgagcatatc 834
Db 1038 ATTCAGATGAAGAAGCTACGAGCCCTGATGGCTGGTTAGATGATGAACCCGGAATATGTA 1097
Qy 835 cctgaccctgatgctaagaagcccgagactgggatgaagagatggagcgagagtgggaa 894

Db 1098 CCTGATCAGATGAGAGAGCCAGAGATTGGATGAAGATATGGATGGAGATGGAG 1157
Qy 895 cccccagtgattcagaaccttgagtacaag-----ggtgagtgaagccc 939
Db 1158 GCTCTCAGATCGCAACCTTAAGTGTGAGTCGCGCCCTGGGTGTGTTGTCGCACGGA 1217
Qy 940 cggcagatcgcacaacccagattacaaggcacttgatccaccacagaattgacaacccc 999
Db 1218 CCTATGATTGACAACCTTAATTATAAGGGCAATGGAAGCCTCCCATGATTGACAATCCT 1277
Qy 1000 gagta 1004
Db 1278 AACTA 1282

RESULT 12
US-08-574-959A-8
; Sequence 8, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3211 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 439..3157
US-08-574-959A-8

Query Match 4.1%; Score 79.6; DB 2; Length 3211;
Best Local Similarity 56.5%; Pred. No. 8.4e-10;
Matches 148; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
Qy 1100 tcataccaacagatgagcgatacctgagggatttggcaacgagacgctggggcgtaacaa 1159
Db 2384 TCAACACAGTGTGATGAGAGGAGGAGGAGGAAGAAGAGAGAGAGAGAGAGAGAGAG 2443
Qy 1160 aggcagcagagaaacaataatgaaggacagagagagagagagagagagagagagag 1219
Db 2444 AAGAAGAGGAG 2503

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 16:47:01 ; Search time 1497.62 seconds
 (without alignments)
 12358.734 Million cell updates/sec

Title: US-09-807-148-7
Perfect score: 1958
Sequence: 1 gaattccggatcatctttaac.....aaaaaaaaaagaattc 1958

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

Db 661 ACCAAGAAGATAAGGATCTGATGCTTCAAAACCGGAAGACTGGGATGAGCGGCCAA 720
QY 780 gatcagatgccacagactccaagctgagactgaggacaagcccgagcatatccctga 839
Db 721 GATCGATGATCCACAGACTCCAAAGCTGAGGACTGGGACAAAGCCCGACATATCCCTGA 780
QY 840 cctgatgctaagaagcccgagactggatgaagagatggagcgagagtgggaaccccc 899
Db 781 CCTGATGCTAAGAAGCCGAGACTGGGATGAAGAGATGGACGAGAGTGGAA-CCCC 839
QY 900 agtgattcagaacctgagtacaaggggtgagtgaagcccgagaccccgagacaaccaga 959
Db 840 AGTGATTCAGAACCTTGAGTACAAAGGCTGAGTGGAGCCCGGACAGTCGACAACCCAGA 899
QY 960 ttacaaggcacttgatcacccagaaattgacaaccccgagatttcccgatcccaag 1019
Db 900 TTACAAGGCACTTGGATCCACCAGAAATTGACAA-CCCGAGTATTCCTCCGATCCCAAG 958
QY 1020 tatctatgctatgataactttggcgctgggctggacctggcgaggtcaagtctgg 1079
Db 959 TATCATGCTTATGATAAATTGGCGTGTGGCGCTGGACCTTGGCAGGTCAAGTCTGG 1018
QY 1080 caccattttgacaacttctcatcaccaacgatgaggcatagcgtgaggtttgg 1136
Db 1019 CACATCCTTTGACAACCTCTCATCACCACGATGARGCATACGCTGAAGAGTTTGG 1075

RESULT 2
AL579227/c 981 bp mRNA EST 16-FEB-2001
LOCUS AL579227 LTI_FL012_Tc1 Homo sapiens cDNA clone CS0DH002YD03 3 prime
DEFINITION mRNA sequence.
ACCESSION AL579227
VERSION AL579227.1 GI:12944062
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 981)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..981
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DH002YD03"
/clone_lib="LTI_FL012_Tc1"
/tissue_type="T cells from T cell leukemia"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
Library was constructed by Life Technologies. Contact :
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 210 a 270 c 256 g 241 t 4 others
ORIGIN

```

Query Match 48.4%; Score 947.6; DB 106; Length 981;
 Best Local Similarity 98.6%; Pred. No. 1.3e-200;
 Matches 974; Conservative 4; Mismatches 3; Indels 7; Gaps 2;

RESULT 3
 AL533447
 LOCUS AL533447 972 bp mRNA EST 13-FEB-2001
 DEFINITION AL533447 LTI_FL015_Brn1 Homo sapiens cDNA clone CS0DN004YJ24 5

```

QY 832 atccctgaccctgatgctaagaagcccgagactggatgaagagatggagcgagagtgg 891
Db 981 ATCCCTFACCCTGATGCTTAAGAAGCCCGAGAACTGGGATGAAGAGATGAACGAGAGT-G 923
QY 892 gaacccccagtgattcagaacccctgagtcacaagggtagtgggaagcccccgagatcgac 951
Db 922 GAACCCCAAGTATTGATTCAGAACCCCTGAGTACAAAGGTTGAGTGAAGCCCGGAGATCGAC 863
QY 952 aacccagattacaaggggcacttgatcccccagaaattgacaacccccgagatttctccc 1011
Db 862 AACCCAGATTACAAAGGGCACTTGGATCCACCCAGAAATTGACAACCCCGAGTATTCTCCC 803
QY 1012 gatcccgatctatgctatgataactttggcgctgggctggacctggacctctgagaggtc 1071
Db 802 GATCCCAAGTATTCTATCCCTATGATAACTTTGGCGTCTGGCCCTGGACCTCTGGCAGGTC 743
QY 1072 aagtctggcaccattcttgacaacttctcatcaccaacgatgagcgatcgtcgtgaggag 1131
Db 742 AAGTCTGGCACCATTCTTGCACAACTTCTCATCACCACGATGAGGCATACGCTGAGGAG 683
QY 1132 ttggcaacgagacgtggggcgtaacaaaggcagcagagagaacaaatgaaggacaacag 1191
Db 682 TTGGCAACGAGACGTTGGGGCGTAAACAAGGCAGCAGAGAAACAAATGAAGGACAAACAG 633
QY 1192 gacgagcagcagaggtcttaaggaggaggaagacaagaacgcaagagggagggag 1251
Db 622 GACGAGCAGCAGAGGCTTAAGGAGGAGGAGAGACACAAGAACGCAAGAGGAGGAGGAG 563
QY 1252 gcagaggacaaggaggatgatgaggacaagatgaggaatgaggaatgagggaggaag 1311
Db 562 GCAGAGGACAAGGAGGATGATGAGGACAAGATGAGGATGAGGAGGATGAGGAGGACAAG 503
QY 1312 gaggaaatgaggaagaatgtcccgccaggccaagagagcgtgtagagagccctg 1371
Db 502 GAGGAAGATGAGGAGGAAGATGTCCCGCCAGGCCAAGACGAGCTGTAGAGAGGCCCTG 443
QY 1372 cctccagggtgactgagcgtcctgcgcagagagcttgcgcgcgaataat 1431
Db 442 CCTCCAGGCTGGACTGAGCGCTGAGCGTCTCTGCCGAGAGCTGCCCGCCCAATAAT 383
QY 1432 gtctctgtgagactcgagaacttctatctttccaggctggttcggatttgggtggat 1491
Db 382 GTCTCTGTGARACTCGAGAACCTTTCATTTTTCAGGCTGTTTCGGATTTCGGGTGGAT 323
QY 1492 ttgtgtttgttccctctccactctccacccctcccgcccttttttttttttttttttt 1551
Db 322 TTGTGTTTGTTCCTCTCTCCTCTCCACTCTCCACCCCTCCCGCCCTTTT-----TTT 269
QY 1552 tttaaaactggatttttatctttgattctctctccttcagcctcaccctcggtctctcttt 1611
Db 268 TTTTAACTGGTATTATTATCTTTGATTCTCTTCAGCCCTCACCCTGGTTCATCTTT 209
QY 1612 ctgtatcaacatcttttctgtcctctgtccctctctctctctctctctctctctctct 1671
Db 208 CTGTGATCAACATCTTTTCTTGCTCTGTCTCCCTTCTCTCATCTCTTAGCTCCCTCAAC 149
QY 1672 ctggggggcagtggttgagagacacagcctgagatttctctctctctctctctctct 1731
Db 148 CTGGGGGGCAGTGGTGTGGAGAAGCMACAGGCGCTGAGATTTTCATCTGCTCTCTTCCCTGG 89
QY 1732 agcccgagagggcagcagaaaggggtggtgtcttccacccccccagcactgaggaaga 1791
Db 88 AGCCCGAGAGGGGAGGAGGAGGGGGTGTCTCTCAACCCCGCCAGCACTGAGGAAGAA 29
QY 1792 cggggctctctcatttccacctcctt 1819
Db 28 CGGGGCTCTCTCATTTTACCCCTCCTCMT 1

```

```
prime, mRNA sequence.
ACCESSION AL533447
VERSION AL533447.1 GI:12796940
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 972)
JOURNAL Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..972
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DN004YJ24"
/clone_lib="LTI_FL015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was constructed by Life Technologies.
Contact : Feng Liang Life Technologies, a division of
Invitrogen 9800 Medical Center Drive Rockville, Maryland
20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 249 a 258 c 269 g 195 t 1 others
ORIGIN

Query Match 47.7% Score 934.8; DB 106; Length 972;
Best Local Similarity 98.9%; Pred. No. 9.1e-198;
Matches 961; Conservative 1; Mismatches 8; Indels 2; Gaps 2;

QY 58 ctgcggaggctctgttttaaaagggccgcgctgtccgcgccctcgg-cgcgcattgctgt 116
|||||
Db 1 CTGCGGAGGTCGTTTAAAGAGGCCGCGCGTTCGCCGCCCTCGGCCGCCATGCTGCT 60
|||||

QY 117 atcgcgtgcgtgtcgtcgcctcctcctcgtgcgcgtgcgcgtgcgcgtgcgtctactt 176
|||||
Db 61 ATCCGTGCCGTGCTGCTGCCCTCCTCGGCTGCGCTGCGCGCGCTGCGCTGCTACTT 120
|||||

QY 177 caagagacagttctgacgagacagcgggtgacttccgcgtggtgacgaatccaacacaa 236
|||||
Db 121 CAAGAGCAGTTCTGGACGAGAGTGGTGACTTCGCCGCTGGATCGAATCCAACACAA 180
|||||

QY 237 gtcagatttgcaaaattcgtctcagttccggca-aggctacggtgacgagagaaag 295
|||||
Db 181 GTCAGATTTTGGCAATTCGTTCTCAGTTCGGCATAGTTCACGGTCACGAGGAGAA 240
|||||

QY 296 ataaagtttgacagaagccaggatgcacgcgttttatgctcgtcgcgcagtttcagc 355
|||||
Db 241 ATAAAGGTTTGCAGAACGCCAGGATGCACGCTTTTATGCTGCTGCGGCCAGTTTCGAGC 300
|||||

QY 356 ctttcacacaaaggccagacgcgtgtgtgagttcagttcagcgtgaaacatagcagaaca 415
|||||
Db 301 CTTTCAGCAACAAGGCCAGACGCTGGTGGTGCAGTTTACCGGTGAAACATGAGCAGACA 360
|||||

QY 416 tcgactgtggggcggtatgtgaagctgtttcctaatagtttgaccagacagacatgc 475
|||||
Db 361 TCGACTGTGGGGCGGCTATGTGAAGCTGTTTCTTAATAGTTTGACACAGACATGC 420
|||||

QY 476 acgagacacgaataacacatcatgtttgtcccgacatcgtggccctggcaccaga 535
|||||
Db 421 ACGGAGACTCAGAATACAAATCATGTGTTGGTCCCGACATCTGTGGCCCTGGCACCAGA 480
|||||
```

Qy	853	aagcccgaggactggatgaagagatggacggagagtgggaaccccccgatgcagaac	912
Db	1030	AAGCCCCAGAGACTGGR-TGAAGAGATGGACGGAGAGTGGGAACCCCGATGATTCAGAAC	972
Qy	913	ccTgagTacaagsggtgagTggaagccccggcagatcgcaaccccgatTacaagggcact	972
Db	971	CCTGAGTACAAGGGTCAGTGGAAAGCCCGCAGATCGACAMCCSAGATTACRAGGCGACT	912
Qy	973	TgataccaccagaataTgaacaaacccgagTattcccgattcccgatcctatgcctat	1033
Db	911	TGATCCACCAGAAATTGACAMCCCGAGATTCTCCCGATCCGAGTATCATGCCAT	852
Qy	1033	gataactttggcgtgctggccctctggaccctctggcaggtcaagTctbpgcaccatctttgac	1093
Db	851	GATAAATTTTGGCTGTGGGCCCTGGACCTCTGGCAGGTCAAGTCTGGCACCATCTTTGAC	792
Qy	1093	aacttctcatcaccaacgatgaggcatacgctgaggagttTggcaacgagacgtTggggc	1153


```
Db 304 GCCCTTTT-----TTTTTTAAACTGGTATTTATCTTTTGATCTCCTTCAGCCCTCAC 251
Qy 1595 ccctgggtctcatctttcttgatacaacatctttcttgctctctgccccttctctcatct 1654
Db 250 CCCTGGTCTCATCTTCTTGATCAACATCTTTCTTGCTCTGTCCCTCTCTCATCT 191
Qy 1655 cttagctccctccaaacctgggggagtggtggtggaagccacagccctgagattcca 1714
Db 190 CTTAGCTCCCTCCAACTGGGGGCGAGTGTGTGGAGAACCCACAGGCCCTGAGATTTCA 131
Qy 1715 tctactctctctgagccagagagagagagagagagagagagagagagagagagagagag 1774
Db 130 TCTGCTCTCTCTCTGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 71
Qy 1775 ccagactgaggaagaaagggtctctctctctctctctctctctctctctctctctctcc 1834
Db 70 CCAGACTGAGAGAGAGGGGCTCTCTCATTTACCCCTTCCCTTCTCCCAAGCCCM 11
Qy 1835 aggactgggc 1844
Db 10 AGGACTGGGC 1

RESULT 11
AL547026 849 bp mRNA EST 16-FEB-2001
LOCUS AL547026 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0DI032YC16 5
DEFINITION prime, mRNA sequence.
ACCESSION AL547026
VERSION AL547026.1 GI:12880714
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 849)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 849
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI032YC16"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site.1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 212 a 222 c 242 g 172 t 1 others
ORIGIN

Query Match 42.08; Score 823; DB 106; Length 849;
Best Local Similarity 99.5%; Pred. No. 7e-173;
Matches 845; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

Qy 60 gccgagaggtcttttaaggg-cgcgcggttgccgccctcgg-cgcgcctgctgcta 117
Db 1 GCCGAGAGGTGCTTTAAAGGGCCGCGGTGGCGCCCTCGGCCCGCATGCTGCTA 60
Qy 118 tccgtgcggtgtgctgcgctcctgcgctggcgctgcgcgagcctgcgctctacttc 177
```

```
Db 61 TCCGTGCGCTGCTGCTCGGCTCCTCGGCTGCGCTGCGGAGCCTGCGCTTACTTC 120
Qy 178 aagcagcagttcttgacgagaggggtggacttcccgctgcatcgaatcaaacacaag 237
Db 121 AAGGAGCAGTTTCTGGAGCAGACGGGTGGACTTCCCGCTGGATCGAATCCAAACACAAG 180
Qy 238 tcagattttggcaaatctcgttctcagttcccggaagttctacggtgacgagagaaagat 297
Db 181 TCAGATTTTGGCAATTCGTTCTCAGTTCGCGCAAGTTCTACGGTGCAGGAGAGAAGAT 240
Qy 298 aaagtttgagacaagccaggatgcacgcttttatgctctgtcggccagtttcgagcct 357
Db 241 AAAGGTTTGCAGACAAGCAGGATGCACGCTTTTATGCTCTGTGCGCCAGTTTCGAGCCT 300
Qy 358 ttcagcaacaagccagacgctggtggtcagttcacggtgaaacatgagcaaacatc 417
Db 301 TTCAGCAACAAGGCCAGACCGCTGTTGTGTCAGTTCACGGTGAACATGACGAGACATC 360
Qy 418 gactgtggggcggtctatgtgaagctgtttcctaatagtttggaccagacagacatgcac 477
Db 361 GACTGTGGGGCGGCTATGTGAAGCTGTTCTTAATAGTTTGGACACAGACATGCAC 420
Qy 478 ggagactcagaatacaacatcatgtttggtcccagacatctgtggccctggcaccagaag 537
Db 421 GGAGACTCAGAATAACAACATCATGTTTGTGCTCCGACATCTGTGGCCCTGGCACCAGAAG 480
Qy 538 gttcatgctcttcaactacaagggcaagacgctgctgcatcacaagagacatccgttgc 597
Db 481 GTTCATGTCTCTTCAACTACAGGGCAAGAACGTGCTGATCAACAAGAGCATCCGTTGC 540
Qy 598 aaggatgatgatttacacacctgtacacactgattgtgcggccagacacacacctatgag 657
Db 541 AAGGATGATGAGTTTACACACCTGTACACACTGATTGTGCGGCCAGACACACCTATGAG 600
Qy 658 gtgaagattgacacagccaggtgagtcggtcctcttggaagcagattggacttcctg 717
Db 601 GTGAAGATTGACAACAGCCAGGTGGAGTCCGGCTCCTTTGGAAGACGATTGGGACTTCCTG 660
Qy 718 ccaccaagaagataaagatccctgatcttcaaacccggaagatggatggagcgggcc 777
Db 661 CCACCAAGAAGATAAAGATGCTGTATGCTTCAAAACCGGAAGACTGGATGAGGCGGCC 720
Qy 778 aagatcgatgatccacagactccaagcctgaggtgaggtggacacccgagcagatccct 837
Db 721 AAGATCGATGATCCACAGACTCCAAGCCTGAGGACTGGACAAAGCCGAGCATATCCCT 780
Qy 838 gacctgatgctaagaagcccgaggtggtgagatgagagatggacgagagtggaaccc 897
Db 781 GACCTGATGCTAAGAAGCCCGAGACTGGGATGAAGAGATGGACGAGAGTGGGAMCCC 840
Qy 898 ccagtgtatt 906
Db 841 CCAGTGATT 849

RESULT 12
AL117505 809 bp mRNA EST 19-OCT-2000
LOCUS AL117505 HEMBA1 Homo sapiens cDNA clone HEMBA1001553 5', mRNA
DEFINITION sequence.
ACCESSION AL117505
VERSION AL117505.1 GI:10932478
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 809)
AUTHORS Oka,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
```

JOURNAL
COMMENT

Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human CDNA project: 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

Location/Qualifiers
1. .809

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBA1001553"
/clone_lib="HEMBA1"
/tissue_type="whole embryo, mainly head"
/dev_stage="embryo, 10 weeks"
/note="vector: pME18SFL3"

BASE COUNT 255 a 180 c 242 g 129 t 3 others
ORIGIN

Query Match 40.9%; Score 800.2; DB 107; Length 809;
Best Local Similarity 99.3%; Pred No. 8.3e-168;
Matches 802; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 527 gcaccaagaagtctatgtcatcttcaactacaagggaagcagctgctgatcaacaagg 586
|||||
Db 2 GCACCAAGAAGGTTTCATGTCATCTTCACTACCAAGGCAAGAACGCTGTGATCAACAAGG 61
QY 587 acatccgttgcaagatgatgatttaacacactgtacacactgtattgtgcggcagaca 646
|||||
Db 62 ACATCCGTTGCAAGATGATGATTTACACACTGTACACACTGTGTCGGCCGACACA 121
QY 647 acacattatgattgaattgacaacagccaggtggagtcgggctccttggagacgatt 706
|||||
Db 122 ACACCTATGAGTGAAGATTGACACAGCCAGGTGGATCCGGCTCCTTGGGAAGACGATT 181
QY 707 gggactctctgcccccaagaagataaaggatcctgtatgcttcaaaaccggaagactggg 766
|||||
Db 182 GGGACTTCTGCGACCCCAAGAAGATAAAGGATCCTGTATGCTTCAAAACCGGAAGACTGGG 241
QY 767 atgacgggccaagatgatccacagactccaagcctgagactgggacacagccggc 826
|||||
Db 242 ATGAGCGGGCCCAAGATGATGATCCACAGACTCCCAAGCCTGAGGACTGGGACAAAGCCCG 301
QY 827 agcatatccctgacccctgatgctaagaagccgaggaactgggatgaagagatggagcagag 886
|||||
Db 302 AGCATATCCCTGACCCCTGATGCTAGAGAGCCCGAGACTGGGATGAAGATGGACGGAG 361
QY 887 agtgggaaccccccagtgattcagaacccctgagtacaagggtgagtgggaagcccccgcaga 946
|||||
Db 362 AGTGGGAACCCCCAGTGTATTCAGACCCCTGAGTACAAGGTTGAGTGAAGCCCGCCGACGA 421
QY 947 tcgacaacccagattacaaggggcaattggatccaccacagaaattgacaaccccgagatt 1006
|||||
Db 422 TCGACAACCCAGATTACAAAGGGCACTTGGATCCACCCAGAAATTGACAACCCCGAGTATT 481
QY 1007 ctcccagatcccatgactatgctctatgataactttggcgctggcgctggacctctggc 1066
|||||
Db 482 CTCCTCGATCCCAAGTATCTATGCTATGATATTAATCTTGGCGTGTCTGGGCTTGGACCTCTGGC 541
QY 1067 aggtcaagtctggcaccatctttgacaacttctctacacacacgatgaggatcacgctg 1126
|||||
Db 542 AGGTCAAGTCTGGCACCATCTTTTGACAACTTCTCATCACCAACCATGAGGACATACGCTG 601
QY 1127 aggaatttggcaacagacgtggggcgtaacaaaggcagcagagagaacaaatgaaggaca 1186
|||||
Db 602 AGGAGTTTGGCAACAGACGTGGGGCGTTAAACAAAGGCGACGAGAGAGAAAATGAAGGACA 661
|||||

QY 1187 aacaggacagagcagagcgttaaggaggaggaagaacaagaacgaacgaagaggag 1246
|||||
Db 662 AACAGGACGAGGAGCAGAGGCTTAAGGAGGAGAGAGACCCAGAACCAAGAGGANG 721
QY 1247 aggaggcagagcacaagaggatgatgaggacaaagatgagatgaggaggatggagg 1306
|||||
Db 722 AGGAGGCGAGGACAAAGGAGGATGATGAGGACAAAGATGANGATGAGGAGGATGANGAGG 781
QY 1307 acaagaggagaagatgaggaggagaagatgt 1334
|||||
Db 782 ACCAGGAGGAAGATGAGGAGGAGGATGT 809
RESULT 13
AL540762
LOCUS AL540762 840 bp mRNA EST 16-FEB-2001
DEFINITION AL540762 LTI_FL002_P11 Homo sapiens cDNA clone CS0DE004YP07 5 prime
, mRNA sequence.
ACCESSION AL540762
VERSION AL540762.1 GI:12871217
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 840)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. .840
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE004YP07"
/clone_lib="LTI_FL002_P11"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 210 a 217 c 241 g 170 t 2 others
ORIGIN
Query Match 40.9%; Score 800; DB 106; Length 840;
Best Local Similarity 99.2%; Pred No. 9.3e-168;
Matches 833; Conservative 2; Mismatches 2; Indels 3; Gaps 3;
QY 58 ctgcggagggtcgctgttaaggagg-cgcggtgtccgcccccctcgcc-gccatgtctgc 115
|||||
Db 1 CTGCGGAGGGTCGTTTAAAGGGCCGCGTYGCGCCGCCCTCGGCTGCCATGCTGC 60
QY 116 tatccgtgcctgtgctgcctcctcgccctgcgcgtgcgcgagcgcgtcctact 175
|||||
Db 61 TATCCGTCGCCGTGCTGTCGCCCTCCTCGGCTCGCCGAGCCTGCCGCTACT 120
QY 176 tcaaggagcagttctgacgagagcgggtgactcccgctggatcgaaatccaacaca 235
|||||
Db 121 TCAAGGACAGTTCTCGACGAGAGCGGTGACTTCCCGCTGGATCGAATCCAACACA 180
QY 236 agtcagatttggcaaatcgttctcagttcccgcaagtctacggtgacgaggagaaag 295
|||||
Db 181 AGTCAGATTTTGGCAAAATTCGTTCTCAGTTTCGGCAAGTTCTACGCTGACGAGGAGAAAG 240
|||||

```
QY 296 ataaagtttgcagacagccagagatgacgcgttttatctgtctgctgcggcagtttcgagc 355
|||||
Db 241 ATAAAGGTTTGCAGACAAGCCAGGATCACGCTTTTATGCTCTGCTCGGCCAGTTTCGAGC 300
|||||
QY 356 ctlttcagacaagaagccagacgctgggtgctgagttccacggtgaaacatgagcagaca 415
|||||
Db 301 CTTTCAGACAACAAGGCCAGACGCTGGTGTGCGAGTTTCACGGGTGAACATGAGCAGAA 360
|||||
QY 416 tcgactgaggggcggtatgtgaagctgtttctctaagtattggaacacagacacatgc 475
|||||
Db 361 TCAGCTGTGGGGCGGCTATGTGNAAGCTGTTCTTAATAGTTTGGACCAGACAGACATGC 420
|||||
QY 476 acggagactcagaatacaacatcatgtttgtcccgacatctgtggccctggcaccaga 535
|||||
Db 421 ACGGAGACTCAGAATACAACATCATGTTGTCCTCGACATCTGTGGCCCTGGCACCAGA 480
|||||
QY 536 aggttcattgctcatcttcacatacaagggaagcaagcgtgctgtgatacaacaagacatcg 595
|||||
Db 481 AGGTTTCATGTCATCTTCAACTACAAAGGGCAAGAACGCTGCTGATCAACAAGGACATCCGTT 540
|||||
QY 596 qcaagatgatgatttacacactgtacacactgattgtgcggccagacacacactatg 655
|||||
Db 541 GCAGGATGATGAGTTTACACACTGTACACACTGATGTGCGGCCAGACAAACACCTATG 600
|||||
QY 656 aggtgaagattgacaacagccaggtgagtcgcggtccttgggaagacgattgggaactcc 715
|||||
Db 601 AGGTGAAGATTGACAACAGCCAGGTGGAGTCCGGCTCCTTGGGAAGACGATTGGGACTTCC 660
|||||
QY 716 tgcacccaagaagataaagatactgctgattcttcaaacccgggaagactgggatgagcggg 775
|||||
Db 661 TGGCACCAAGAAGATAAAGGATCCTGATGCTTCAAAACCGGAAGACTGGGATGAGCGGG 720
|||||
QY 776 ccaagatcgatgatcccaagactccaagctgaggaactgggaacagccagacatctcc 835
|||||
Db 721 CCAAGATCGATGATGCCACAGACTCCAAGGCTGAGGACTGGGACAAAGCCGAGCATATCC 780
|||||
QY 836 ctgaccttg-atgctaagaagccgagactgggataagagatggacgagagatgggaa 894
|||||
Db 781 CTGACCTTGATGCTAAGAGCCGAGGACTGGGATGAAGAGATGGACGGRGAGTGGGGA 840
|||||

RESULT 14
BG491927/c
LOCUS
DEFINITION BG491927 956 bp mRNA EST 27-MAR-2001
602535842T1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684763 3',
mRNA sequence.
ACCESSION BG491927
VERSION BG491927.1 GI:13453439
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM1493 row: d column: 12
High quality sequence start: 3
High quality sequence stop: 911.
Location/Qualifiers
1..956
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:4684763"
/tissue_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

BASE COUNT 168 a 269 c 235 g 284 t
ORIGIN

```
Query Match 40.5%; Score 795.8; DB 154; Length 956;
Best Local Similarity 96.5%; Pred. No. 8.2e-167;
Matches 856; Conservative 0; Mismatches 27; Indels 4; Gaps 4;

QY 397 gtgaacatagcagaacatcgactgtggggcggtctatgtgaagctgtttccctaagt 456
|||||
Db 942 GGGGAACAAGAGCAGACATCGACTGT-GGGCGGCTATGTGAAGCTGTTT-CTAATAGT 885
|||||
QY 457 ttgacacagacagacatgcacggagactcagaatcacacat-catgtttggtcccgacat 515
|||||
Db 884 TGGACCCAGACAGACATGCACGGAGACTCAGGATACAACATCCATGTTTGGTCCCGACAT 825
|||||
QY 516 ctgtggccctggcaccagaaggttcatgtcatcttcaactacaaagggcaa-gaacgtgc 574
|||||
Db 824 CTGTGGCCCTGGCACCAAGAAGGTTTCATGTCATCTTCAACTACAAGGSCAAGGAACGTGC 765
|||||
QY 575 tgatcaacaggagacatccgttgcaaggatgatgagtttacacactgtacacactgattg 634
|||||
Db 764 TGATCAACAAGAGACATCCGTTGCAAGGATGATGAGTTTACACACTGTACACACTGATTG 705
|||||
QY 635 tgcggccagacaacacactatgagtgagattgacaacagccaggtgagtcgggtccct 694
|||||
Db 704 TCGGCCACAGACAACACTATGAGTGAAGATTGACACAGCCAGGTGGAGTCCGGCTCCT 645
|||||
QY 695 tggaagacgattgggacttctccaccaccaagaagataaaaggatcctgatgtctcaaac 754
|||||
Db 644 TGAAGACGATTGGGACTTCTCTGCCACCAAGAAGATAAAGGATCCTGATCTTCAAAAC 585
|||||
QY 755 cgaagactgggtgagcgggccaagatcgatgatccacagactccaagctgaggact 814
|||||
Db 584 CGGAAGACTGGGATGAGCGGGCCAAAGATCGATGATGCCACAGACTCCAAGCCTGAGGACT 525
|||||
QY 815 gggacaagcccgagcatatccctgacctgatgctaagaagcccgagactgggatgaag 874
|||||
Db 524 GGGACAAGCCGGAGCATATCCTGTACCCTGATGCTAGAGAGCCCGAGACTGGGATGAAG 465
|||||
QY 875 agatggcagagagtggaacccccagtgattcagaacccctgagtacaagggtagtga 934
|||||
Db 464 AGATGGACGGAGAGTGGGAACCCCACTGATTTCAAGACCCTGAGTACAAGGGTGAAGTGA 405
|||||
QY 935 agccccggcagatgcacaacccagattacaaggcacttggatccaccagaaattgaca 994
|||||
Db 404 AGCCCCGGCAGATCGACAACCCAGATTACAAGGGCACCTTGGATCCACCCAGAAATTGACA 345
|||||
QY 995 accccgagttatcccgactccagttatctatgctctgataaactttgggtgctgggccc 1054
|||||
Db 344 ACCCCGAGTATTTCCCGATCCCACTATCTATGCTATGATAACTTTGGCGTGTGGGCC 285
|||||
QY 1055 tggacctctggcaggtcaagcttggcaccatctttgacaactctctcatcaccacacgatg 1114
|||||
Db 284 TGGACCTCTGGCAGGTCAAGTCTGGCACCCTCTTTGACAACCTTCTCTCATCAACAACATG 225
|||||
QY 1115 aggcatacgtgagagatttggcaacagacgctggggcgctaaacaaaggcagcagagaac 1174
|||||
Db 224 AGGCATACGTGAGGAGTTTGGCAACGAGAGCTGGGGCGTAAACAAGGACGACGAGAGAAC 165
|||||
```


Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	177	55.7	415	4	US-08-675-816-2		Sequence 2, Appli
2	115	36.2	593	1	US-08-296-362-2		Sequence 2, Appli
3	79	24.8	542	4	US-08-673-816-6		Sequence 6, Appli
4	61.5	19.3	582	3	US-08-908-865-3		Sequence 3, Appli
5	54.5	17.1	410	2	US-08-723-415B-10		Sequence 10, Appli
6	54.5	17.1	410	4	US-09-189-627A-10		Sequence 10, Appli
7	54	17.0	405	1	US-07-829-954-2		Sequence 2, Appli
8	54	17.0	405	1	US-07-994-423-2		Sequence 2, Appli
9	54	17.0	405	1	US-08-421-891-2		Sequence 2, Appli
10	54	17.0	816	1	US-08-190-802A-54		Sequence 54, Appli
11	53.5	16.8	279	2	US-08-701-131A-23		Sequence 23, Appli
12	53.5	16.8	355	1	US-08-293-549-6		Sequence 6, Appli
13	53.5	16.8	812	1	US-08-446-794A-4		Sequence 4, Appli
14	53.5	16.8	984	2	US-08-673-789-9		Sequence 9, Appli
15	53.5	16.8	984	2	US-08-449-645A-19		Sequence 19, Appli
16	53.5	16.8	984	2	US-08-702-367A-19		Sequence 19, Appli
17	53.5	16.8	984	5	PCR-US95-04681-19		Sequence 19, Appli
18	53	16.7	410	2	US-08-723-415B-11		Sequence 11, Appli
19	53	16.7	410	2	US-08-428-131-2		Sequence 2, Appli
20	53	16.7	410	2	US-08-603-846-2		Sequence 2, Appli
21	53	16.7	410	4	US-09-078-596-2		Sequence 2, Appli
22	53	16.7	410	4	US-09-189-627A-11		Sequence 11, Appli
23	52.5	16.5	236	1	US-08-307-499-28		Sequence 28, Appli
24	52.5	16.5	236	4	US-09-299-268-28		Sequence 28, Appli
25	52.5	16.5	329	4	US-09-036-987A-26		Sequence 26, Appli
26	52.5	16.5	770	1	US-08-525-654A-1		Sequence 1, Appli
27	52.5	16.5	771	1	US-08-525-654A-3		Sequence 3, Appli

```
RESULT 2
US-08-296-362-2
; Sequence 2, Application US/08296362
; Patent No. 5691306
; GENERAL INFORMATION:
; APPLICANT: Bergeron, John J.M.
; APPLICANT: Thomas, David Y.
; APPLICANT: Wada, Ikuro
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; TITLE OF INVENTION: PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
; TITLE OF INVENTION: PROTEIN PRODUCTION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,362
; CLASSIFICATION: 435
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Deehr, Manya S.
; REGISTRATION NUMBER: 37,120
; REFERENCE/DOCKET NUMBER: 690066.401C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-296-362-2

Query Match 36.2%; Score 115; DB 1; Length 593;
Best Local Similarity 37.5%; Pred. No. 1.3e-07;
Matches 24; Conservative 12; Mismatches 18; Indels 10; Gaps 1;

QY 6 KHVIFNYGKNNVINKDIRCK-----DDEFTHTLTLVLRPDNTYEVKIDNSQVE 55
||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 200 KLFIFERHKNPKGVYERHAKRPDADLKYFTDKKTHLYTLILNPDNSFEILVDQSIVN 259

QY 56 SGSL 59
||:|
Db 260 SGNL 263

RESULT 3
US-08-675-816-6
; Sequence 6, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND USES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
```

```
ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206-622-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-675-816-6

Query Match 24.8%; Score 79; DB 4; Length 542;
Best Local Similarity 30.2%; Pred. No. 0.01;
Matches 19; Conservative 10; Mismatches 24; Indels 10; Gaps 3;

QY 1 GP----GTRKKVHVIFNYKG--KNVLINKDIR----CKDDEFTHTLTLVLRPDNTYEVKID 50
||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 143 GPDKCGATNKVHFILKHUKPKSGEYIEHLKYPSPVPSDKLTHVYTAILKPDNELRLVD 202

QY 51 NSQ 53
;
Db 203 GEE 205

RESULT 4
US-08-906-865-3
; Sequence 3, Application US/08906865
; Patent No. 6040168
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; APPLICANT: Porton, Barbara
; APPLICANT: Kao, Hung-Ten
; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,865
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
```

TELEFAX: 201-343-1684
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 582 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: /desc = "Synapsin IIa"
 HYPOTHEICAL: NO
 ORGANISM: Homo sapiens
 US-08-906-865-3

Query Match 19.3%; Score 61.5; DB 3; Length 582;
 Best Local Similarity 40.0%; Pred. No. 2.7;
 Matches 14; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

QY 13 YKGNVLNLIKDKDDFTHLYTLVIRPDNTYEV 47
 DB 130 FRGKKVLGDYDKVEQAEFSEL-NLVAHADGTAV 163

RESULT 5

US-08-723-415B-10
 ; Sequence 10, Application US/08723415B
 ; Patent No. 5859199
 ; GENERAL INFORMATION:
 ; APPLICANT: LaThangue, Nicholas B.
 ; APPLICANT: de la Luna, Susana
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHYE P.C.
 ; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
 ; CITY: Arlington
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22201-4741
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/723,415B
 ; FILING DATE: 30-SEP-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9610195.1
 ; FILING DATE: 15-MAY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crawford, Arthur R.
 ; REGISTRATION NUMBER: 25,327
 ; REFERENCE/DOCKET NUMBER: 117-220
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-816-4000
 ; TELEFAX: 703-816-4100
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 410 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-723-415B-10

Query Match 17.1%; Score 54.5; DB 2; Length 410;
 Best Local Similarity 28.6%; Pred. No. 16;
 Matches 20; Conservative 11; Mismatches 18; Indels 21; Gaps 5;

QY 2 PGTKKVH---VIFNYKGNVLNLIKDKDDFTHLYTLVIRPDNTYEVKIDNSQV----- 54
 DB 252 PPSVHLPFIIVNTSKTKVI---DCSISNDKFEYLFNF-----DNTFEIH-DDIEVLKRM 303
 QY 55 -----ESGS 58
 DB 304 GMACGLESGS 313

RESULT 6

US-09-189-627A-10
 ; Sequence 10, Application US/09189627A
 ; Patent No. 6159691
 ; GENERAL INFORMATION:
 ; APPLICANT: La Thangue, Nicholas
 ; APPLICANT: de la Luna, Susana
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
 ; FILE REFERENCE: 620-54
 ; CURRENT APPLICATION NUMBER: US/09/189,627A
 ; CURRENT FILING DATE: 1998-11-10
 ; PRIOR APPLICATION NUMBER: 08/723,415
 ; PRIOR FILING DATE: 1996-09-30
 ; PRIOR APPLICATION NUMBER: GB 9610195
 ; PRIOR FILING DATE: 1996-05-15
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 410
 ; TYPE: PRT
 ; ORGANISM: human
 ; US-09-189-627A-10

Query Match 17.1%; Score 54.5; DB 4; Length 410;
 Best Local Similarity 28.6%; Pred. No. 16;
 Matches 20; Conservative 11; Mismatches 18; Indels 21; Gaps 5;

QY 2 PGTKKVH---VIFNYKGNVLNLIKDKDDFTHLYTLVIRPDNTYEVKIDNSQV----- 54
 DB 252 PPSVHLPFIIVNTSKTKVI---DCSISNDKFEYLFNF-----DNTFEIH-DDIEVLKRM 303
 QY 55 -----ESGS 58
 DB 304 GMACGLESGS 313

RESULT 7

US-07-329-954-2
 ; Sequence 2, Application US/07829954
 ; Patent No. 5344819
 ; GENERAL INFORMATION:
 ; APPLICANT: HAMMOND, Geoffrey L.
 ; TITLE OF INVENTION: Pharmaceutical Composition For Use In
 ; TITLE OF INVENTION: Treating Inflammation
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/829,954
 ; FILING DATE: 19920203
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/664,114

10


```

; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,191A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/088
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-701-191A-23

Query Match          16.8%; Score 53.5; DB 2; Length 279;
Best Local Similarity 36.4%; Pred. No. 13;
Matches 12; Conservative 9; Mismatches 11; Indels

QY      16 KNLVINKDIRKDDFEHLYLIVRPDNTVEVK 48
       :|:::| | | | | | | | | |
Db      143 RNILVNQLCKKVSDF-GLTRLLDDFDGTVETQ 174

RESULT 12
US-08-292-549-6
; Sequence 6, Application US/08292549
; Patent No. 5464938

```

APPLICANT: Smith, Craig A.
 APPLICANT: Goodwin, Raymond G.
 TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/292,549
 FILING DATE:
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/963,330
 FILING DATE: 10/19/92
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia A.
 REGISTRATION NUMBER: 34,693
 REFERENCE/DOCKET NUMBER: 2602-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 587-0430
 TELEFAX: (206) 233-0644
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 355 amino acids

COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,549
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/963,330
FILING DATE: 10/19/92
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2602-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-292-549-6

Query Match 16.8%; Score 53.5; DB 1; Length 355;
Best Local Similarity 26.0%; Pred. No. 18;
Matches 13; Conservative 12; Mismatches 16; Indels 9; Gaps 2;
QY 18 VLINKDI-----RCKDEF-THLYTLIVRPDNTYEVKIDNSQVESGS 58
Db 14 IILNSDITPEPSNGKCKDNEYKRHHLCCLSCPPGTYSRLCDSKTNTNT 63

RESULT 13

US-08-446-794A-4
Sequence 4, Application US/08446794A
Patent No. 5747327
GENERAL INFORMATION:
APPLICANT: UEKI, JUN
APPLICANT: MORIOKA, SHINJI
TITLE OF INVENTION: PHOSPHOLIPASE D GENE ORIGINATED FROM
TITLE OF INVENTION: PLANT
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: P.O. BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,794A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 0760-0203P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 4:
LENGTH: 812 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-794A-4

Query Match 16.8%; Score 53.5; DB 1; Length 812;
Best Local Similarity 26.3%; Pred. No. 52;
Matches 15; Conservative 11; Mismatches 26; Indels 5; Gaps 1;

QY 1 GPGTKKHVIFNYKGNV-----LINKDIRCKDDETHLYTLIVRPDNTYEVKIDNS 52
Db 48 GRGATKIYATVLEKARVGRTMISNPVPRWYEFHYCAHMAADVFTVKIDNS 104

RESULT 14

US-08-673-789-9
Sequence 9, Application US/08673789
Patent No. 5814479
GENERAL INFORMATION:
APPLICANT: ZHOU, REMPING; SCHULZ, NICHOLAS;
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,

APPLICANT: GEORGE, F.
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673,789
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/177,812
FILING DATE: 04-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 984
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-673-789-9

Query Match 16.8%; Score 53.5; DB 2; Length 984;
Best Local Similarity 36.4%; Pred. No. 67;
Matches 12; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 16 KNVLINKDIRCKDDETHLYTLIVRPDNTYEVK 48
Db 761 RNLVNONLCKVSDP-GLTRLDDDFDGTETQ 792

RESULT 15

US-08-449-645A-19
Sequence 19, Application US/08449645A
Patent No. 5981245
GENERAL INFORMATION:

APPLICANT: FOX, GARY M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 984 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-449-645A-19

```

```

Query Match      16.8%; Score 53.5; DB 2; Length 984;
Best Local Similarity 36.4%; Pred. No. 67;
Matches 12; Conservative 9; Mismatches 11; Indels 1; Gaps 1;
Qy 16 KNVLKDIRCKDDEFTILYTLIVRPDNTYEVK 48
Db 761 RNILVNQLCKKVSDF-GLTRLLDDFDGTYETQ 792

```

Search completed: October 21, 2001, 03:06:42
Job time: 332 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:07:44 ; Search time 52.92 Seconds
(without alignments)
86.366 Million cell updates/sec

Title: US-09-807-148-8
 Perfect score: 318
 Sequence: 1 GPCTKKVHVIFNYKGNVLI.....PDNTYEVKIDNSOVESGSL 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	318	100.0	416	1	S06763	calreticulin precu	
2	318	100.0	416	2	JH0819	calreticulin precu	
3	318	100.0	417	1	A37047	calreticulin precu	
4	318	100.0	418	1	A34154	calreticulin precu	
5	313	98.4	400	2	S43376	calreticulin, brai	
6	313	98.4	421	2	S36799	calreticulin precu	
7	281	88.4	405	1	JH0795	calreticulin precu	
8	278	87.4	384	2	S29130	calreticulin (clon	
9	278	87.4	411	2	S29129	calreticulin precu	
10	269	84.6	406	2	A56637	calreticulin homol	
11	262	82.4	419	2	S71343	calreticulin precu	
12	242	76.1	336	2	A32507	41K larval antigen	
13	237	71.4	393	1	A48573	calreticulin autoa	
14	210.5	66.2	395	2	S25851	calreticulin precu	
15	184	57.9	416	2	T14554	calreticulin - bee	
16	179	56.3	421	2	S58170	calreticulin precu	
17	177	55.7	415	2	T10172	calreticulin - cas	
18	176	55.3	412	2	T05703	calreticulin - bar	
19	176	55.3	415	2	T05705	calreticulin - bar	
20	172	54.1	425	2	C96605	calreticulin (Ctrl	
21	169	53.1	389	2	T03691	calreticulin - com	
22	169	53.1	416	2	T16968	calreticulin call	
23	169	53.1	444	2	H86224	hypothetical prote	
24	168	52.8	442	2	T07841	probable calreticu	
25	117	36.8	592	2	I53260	calnexin - human	
26	117	36.8	592	2	A46673	calnexin precursor	
27	116	36.5	591	2	B54354	calnexin precursor	
28	116	36.5	591	2	C54354	calnexin precursor	
29	115	36.2	593	1	A37273	calnexin precursor	

ALIGNMENTS

RESULT 1

calreticulin precursor - mouse
S06763.
N:Alternate names: 5K calcium-binding reticuloplasmin; calregulin
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10
C:Accession: S06763; JCI144; PC1233; A57498
R:Smith, M.J.; Koch, G.L.E.

K:Smith, M.J.; Koch, G.L.E.:
EMBO J. 8, 3581-3586, 1989
A:Title: Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a
A:Reference number: S06763; MUID:90059955

A;Residues: 1-416 <SMI>
A;Cross-references: EMBL:X14926; NTD:g50567; PIDN:CAA3053.1; PID:g50568
P;Mazzarella, R.A.; Gold, P.; Cunningham, M.; Green, M.

Gene 120, 217-225, 1992
A;Title: Determination of the sequence of an expressible cDNA clone encoding Erp60/ca
A;Reference number: JCI444; MUID:93013037

A;Molecule type: mRNA
A;Residues: 1-416 <NAZ>
A;Cross-references: GB:M92988; NID:g193084; PIDN:AAA37569.1; PID:g193085
A;Accession: PC1233

A; Molecule type: protein
A; Residues: 18-41 <MA2>
R; White, T.K.; Zhu, Q.; Tanzer, M.L.
J. Biol. Chem. 270, 15926-15929, 1995

J. Biol. Chem. 270, 13928-13929, 1995

A;Accession: A57498
A;Status: preliminary
A;Molecule type: protein
A;Residues: 74-80;142-151;186-193 <WHI>

C;Superfamily: calreticulin
C;Keywords: calcium binding
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-416/Product: calregulin #status experimental <MAY>
F;413-416/Region: endoplasmic reticulum retention signal
C;Superfamily: calreticulin

Query Match 100.0%; Score 318; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. NO. 4.9e-30;
Matches 60; Conservative 0; Mismatches 0; Indels

Qy 1 GPCTKKVHVIENYGKGNVLNKDIRCKDDEFTHLYTLIVRPDNTVEVKIDNSQVESGSL 60
 |||||
 Db 138 GPCTKKVHVIENYGKGNVLNKDIRCKDDEFTHLYTLIVRPDNTVEVKIDNSQVESGSL 197

RESULT
JH0819
2

calreticulin precursor - rat
N:Alternate names: calcium-binding protein 3
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence revision 20-Aug-1994 #text change 20-Jun-2000
C:Accession: JH0819; A49176; S11205; PCL109; S45036; S04867; S39372; A34473; S13045
R:Nakamura, M.; Moriya, M.; Baba, T.; Michikawa, Y.; Yamamoto, T.; Arai, K.; Okinaga, S.
Exp. Cell Res. 205, 101-110, 1993
A:Title: An endoplasmic reticulum protein, calreticulin, is transported into the acrosome
A:Reference number: A49176; MUID:93202172
A:Accession: JH0819
A:Molecule type: mRNA
A:Residues: 1-416 <NAK>
A:CROSS-references: GB:D78308; NID:g1089798; PIDN:BAAL1345.1; PID:g1845572
A:Accession: A49176
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-416 <NAK>
A:CROSS-references: GB:D78308; NID:g1089798; PIDN:BAAL1345.1; PID:g1845572
A:Experimental source: Sprague-Dawley, spermatogenic cells
A:Note: sequence extracted from NCBI backbone (NCBIN:127639, NCBIP:127643)
R:Murthy, K.K.; Banville, D.; Srikanth, C.B.; Carrier, F.; Holmes, C.; Bell, A.; Patel, Y.
Nucleic Acids Res. 18, 4933, 1990
A:Title: Structural homology between the rat calreticulin gene product and the Onchocerca
A:Reference number: S11205; MUID:90370496
A:Accession: S11205
A:Molecule type: mRNA
A:Residues: 1-416 <NAK>
A:CROSS-references: EMBL:X53363; NID:g55854; PIDN:CAA37446.1; PID:g55855
R:Nakamura, M.; Michikawa, Y.; Baba, T.; Okinaga, S.; Arai, K.
Biochem. Biophys. Res. Commun. 186, 668-673, 1992
A:Title: Calreticulin is present in the acrosome of spermatozoa of rat testis.
A:Reference number: PCL109; MUID:92360010
A:Accession: PCL109
A:Molecule type: protein
A:Residues: 18-32 <NAK>
A:Experimental source: testis, strain Sprague-Dawley
R:Soennichsen, B.; Fuellekrug, J.; van Nguyen, P.; Diekmann, W.; Robinson, D.G.; Mieskes
submitted to the EMBL Data Library, May 1994
A:Description: Retention and retrieval: both mechanisms cooperate to maintain calreticulin
A:Reference number: S45036
A:Accession: S45036
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <SOE>
A:CROSS-references: EMBL:X79327; NID:g488840; PIDN:CAA55890.1; PID:g488841
R:Lone, Y.C.; Bailly, A.; Latruffe, N.
submitted to the EMBL Data Library, December 1988
A:Reference number: S04867
A:Accession: S04867
A:Molecule type: mRNA
A:Residues: 'R', 270-358, 'AAG' <LON>
A:CROSS-references: EMBL:X13702; NID:g56055; PIDN:CAA31987.1; PID:g930260
A:Note: the authors designated the protein as D-beta-hydroxybutyrate dehydrogenase
R:Yokoi, T.; Nagayama, S.; Kajiwara, R.; Kawaguchi, Y.; Horiuchi, R.; Kamataki, T.
Biochim. Biophys. Acta 1158, 339-344, 1993
A:Title: Identification of protein disulfide isomerase and calreticulin as autoimmunity
A:Reference number: S39371; MUID:94072621
A:Accession: S39371
A:Molecule type: protein
A:Residues: 18-23, X', 25-32 <YOK>
R:van, P.N.; Peter, F.; Soelling, H.D.
J. Biol. Chem. 264, 17494-17501, 1989
A:Title: Four intracisternal calcium-binding glycoproteins from rat liver microsomes with
active calcium sequestering rat liver vesicles.
A:Reference number: A34473; MUID:90008920
A:Accession: A34473
A>Status: preliminary
A:Molecule type: protein
A:Residues: 18-36 <VAN>
R:Trevies, S.; de Mattel, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld
Biochem. J. 271, 473-480, 1990
A:Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage
A:Reference number: S13045; MUID:91054414

A:Accession: S13045
A:Molecule type: protein
A:Residues: 18-29 <TRE>
C:Superfamily: calreticulin
C:Keywords: calcium binding; glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-416/Product: calreticulin #status experimental <MAT>
F:204-212/Region: nuclear location signal
F:413-416/Region: endoplasmic reticulum retention signal
F:344/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 100.0%; Score 318; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 4.9e-30;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GPQTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 60
DB 138 GPQTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 197
RESULT 3
A37047
calreticulin precursor - human
N:Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding pr
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 18-Feb-2000
C:Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075
R:MCCauliffe, D.P.; Yang, Y.S.; Wilson, J.; Sontheimer, R.D.; Capra, J.D.
J. Biol. Chem. 267, 2557-2562, 1992
A:Title: The 5'-flanking region of the human calreticulin gene shares homology with t
A:Reference number: A42330; MUID:92129342
A:Accession: A42330
A:Molecule type: DNA
A:Residues: 1-417 <MC2>
A:Note: sequence extracted from NCBI backbone (NCBIN:78537, NCBIP:78536)
R:MCCauliffe, D.P.; Lux, F.A.; Lieu, T.S.; Sanz, I.; Hanke, J.; Newkirk, M.M.; Bachin
J. Clin. Invest. 85, 1379-1391, 1990
A:Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/
A:Reference number: A37047; MUID:90237213
A:Accession: A37047
A:Molecule type: mRNA
A:Residues: 1-417 <MCC>
A:CROSS-references: GB:M32294; NID:g337486; PIDN:AAA36582.1; PID:g337487
A:Note: the authors translated the codon GTA for residue 349 as Tyr
R:Rocheach, L.A.; Haselby, J.A.; Meilof, J.F.; Smeenk, R.J.; Unnasch, T.R.; Greene, B.
J. Immunol. 147, 3031-3039, 1991
A:Title: Characterization of the autoantigen calreticulin.
A:Reference number: A46452; MUID:92013129
A:Accession: A46452
A:Molecule type: mRNA
A:Residues: 1-417 <ROK>
A:CROSS-references: GB:M4739; NID:g179881; PIDN:AAA51916.1; PID:g179882
A:Note: sequence extracted from NCBI backbone (NCBIN:60749, NCBIP:60750)
R:Lieu, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.
J. Clin. Invest. 82, 96-101, 1988
A:Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence
A:Reference number: A28812; MUID:88273610
A:Accession: A28812
A:Molecule type: protein
A:Residues: 18-41 <LIE>
A:Note: 18-Ala was also found
R:Dupaix, M.; Schaefer, E.; Krause, K.H.; Tschopp, J.
J. Exp. Med. 177, 1-7, 1993
A:Title: The calcium-binding protein calreticulin is a major constituent of lytic gra
A:Reference number: PH1525; MUID:93115648
A:Accession: PH1525
A:Molecule type: protein
A:Residues: 18-27 <DUP>
A:Experimental source: LAK cell
R:Roijani, M.V.; Finlay, B.B.; Gray, V.; Dedhar, S.
Biochemistry 30, 9859-9866, 1991
A:Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (c

A:Reference number: A40346; MUID:92002034
A:Accession: A40346
A:Molecule type: protein
A:Residues: 18-34; R' <RQ>
R:Krause, K.H.; Simmerman, H.K.B.; Jones, L.R.; Campbell, K.P.
Biochem. J. 270, 545-548, 1990
A:Title: Sequence similarity of calreticulin with a Ca(2+)-binding protein that co-purified with the Golgi apparatus
A:Reference number: S11475; MUID:90380058
A:Accession: S11475
A:Molecule type: protein
A:Residues: 18-32 <KRA>
R:Lamerdin, J.; McCrea, P.; Stillwagen, S.; Ramirez, M.; Carrano, A.
submitted to the EMBL Data Library, November 1996
A:Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region
A:Reference number: Z2906
A:Accession: T45075
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-417 <LAM>
A:Cross-references: EMBL:AD000092; PIDN:AA851176.1
A:Experimental source: cell line 5HL2-8; fibroblast
C:Comment: Autoantibodies specific for this protein are found in Sjogren's syndrome and C:Genetics:
A:Gene: GDB:CALR
A:Cross-references: GDB:125179; OMIM:109091
A:Map position: 19p13.3-19p13.2
A:Introns: 31/1; 65/1; 133/1; 164/3; 234/3; 272/3; 320/3; 351/3
A:Note: CRRC
C:Superfamily: calreticulin
C:Keywords: calcium binding; integrin binding
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-417/Product: calreticulin #status predicted <MAT>
F:414-417/Region: endoplasmic reticulum retention signal

Query Match 100.0%; Score 318; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.9e-30;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
|||||
Db 138 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 197
|||||

RESULT 4
A34154
calreticulin precursor, skeletal muscle - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34154; S13047
R:Fliegel, L.; Burns, K.; MacLennan, D.H.; Reithmeier, R.A.F.; Michalak, M.
J. Biol. Chem. 264, 21522-21528, 1989
A:Title: Molecular cloning of the high affinity calcium-binding protein (calreticulin) cDNA
A:Reference number: A34154; MUID:90094320
A:Accession: A34154
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <FLI>
A:Cross-references: GB:J05138; NID:gi64858; PIDN:AAA31188.1; PID:gi64859
R:Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld
Biochem. J. 271, 473-480, 1990
A:Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage
A:Reference number: S13045; MUID:91054414
A:Accession: S13047
A:Molecule type: protein
A:Residues: 19-32 <TRE>
C:Superfamily: calreticulin
C:Keywords: skeletal muscle
F:1-17/Domain: signal sequence #status predicted <SIG>
F:415-418/Region: endoplasmic reticulum retention signal

Query Match 100.0%; Score 318; DB 1; Length 418;

Best Local Similarity 100.0%; Pred. No. 4.9e-30;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
|||||
Db 138 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 197
|||||

RESULT 5
S43376
calreticulin, brain isoform 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Oct-1994 #sequence_revision 23-Mar-1995 #text_change 07-May-1999
C:Accession: S43376; S36801
R:Matsuoka, K.; Seta, K.; Yamakawa, Y.; Okuyama, T.; Shinoda, T.; Isobe, T.
Biochem. J. 298, 435-442, 1994
A:Title: Covalent structure of bovine brain calreticulin.
A:Reference number: S43376; MUID:94183174
A:Accession: S43376
A:Molecule type: protein
A:Residues: 1-400 <MAT>
A:Experimental source: brain
R:Liu, N.; Fine, R.E.; Johnson, R.J.
Biochim. Biophys. Acta 1202, 70-76, 1993
A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin
A:Reference number: S36799; MUID:93385184
A:Accession: S36801
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 45-63, 'E', 65-83 <LIU>
A:Experimental source: brain, clone 8.1
C:Superfamily: calreticulin
C:Keywords: calcium binding; glycoprotein
F:397-400/Region: endoplasmic reticulum retention signal
F:120-146/Disulfide bonds: #status experimental
F:162/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 98.4%; Score 313; DB 2; Length 400;
Best Local Similarity 98.3%; Pred. No. 1.8e-29;
Matches 59; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
|||||
Db 121 GPCTKKVHVIFNYGKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
|||||

RESULT 6
S36799
calreticulin precursor, brain isoform 2 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Dec-1993 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999
C:Accession: S36799; S36800
R:Liu, N.; Fine, R.E.; Johnson, R.J.
Biochim. Biophys. Acta 1202, 70-76, 1993
A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin
A:Reference number: S36799; MUID:93385184
A:Accession: S36799
A:Molecule type: mRNA
A:Residues: 1-421 <LIU>
A:Cross-references: GB:L13462; NID:g348693; PIDN:AA37307.1; PID:g348694
A:Experimental source: brain, clone 9.4
A:Accession: S36800
A:Molecule type: protein
A:Residues: 35-45 <LI2>
C:Superfamily: calreticulin
C:Keywords: calcium binding; glycoprotein
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-421/Product: calreticulin, brain isoform 2 #status predicted <MAT>
F:418-421/Region: endoplasmic reticulum retention signal
F:141-167/Disulfide bonds: #status predicted
F:183/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.4%; Score 313; DB 2; Length 421;
Best Local Similarity 98.3%; Pred. No. 1.9e-29;
Matches 59; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 60
|||||
Db 142 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 201
|||||

RESULT 7
JH0795
calreticulin precursor - California sea hare
N:Alternate names: protein 407
C:Species: Aplysia californica (California sea hare)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0795; B31409; F60977
R:Kennedy, T.E.; Kuhl, D.; Barzilai, A.; Sweatt, J.D.; Kandel, E.R.
Neuron 9, 1013-1024, 1992.
A:Title: Long-term sensitization training in aplysia leads to an increase in calreticulin
A:Reference number: JH0795; MUID:93098937
A:Accession: JH0795
A:Molecule type: mRNA
A:Residues: 1-405 <KEN>
A:Cross-references: GB:S51239; NID:g262053; PIDN:AAB24569.1; PID:g262054
A:Experimental source: abdominal ganglion and antral nervous system
R:Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.; Sweatt, J.D.
Proc. Natl. Acad. Sci. U.S.A. 85, 7008-7012, 1988
A:Title: Sequencing of proteins from two-dimensional gels by using in situ digestion and
tion in Aplysia.
A:Reference number: A94207; MUID:88320566
A:Accession: B31409
A:Molecule type: protein
A:Residues: 'X', 17-28, 'X', 30-31 <KE2>
R:Sweatt, J.D.; Kennedy, T.E.; Wager-Smith, K.; Gawinowicz, M.A.; Barzilai, A.; Karl, K.
Electrophoresis 10, 152-157, 1989
A:Title: Development of a database of amino acid sequences for proteins identified and
A:Reference number: A60977; MUID:89276264
A:Accession: F60977
A:Molecule type: protein
A:Residues: 'X', 17-28, 'X', 30-31 <SWE>
C:Superfamily: calreticulin
C:Keywords: calcium binding; endoplasmic reticulum
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-405/Product: calreticulin #status experimental <MAT>
F:402-405/Region: endoplasmic reticulum retention signal

Query Match 88.4%; Score 281; DB 1; Length 405;
Best Local Similarity 85.0%; Pred. No. 1.2e-25;
Matches 51; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 60
|||||
Db 134 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 193
|||||

RESULT 8
S29130
calreticulin (clone 8) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: S29130; T01068
R:Treves, S.; Zorzato, F.; Pozzan, T.
Biochem. J. 287, 579-581, 1992
A:Title: Identification of calreticulin isoforms in the central nervous system.
A:Reference number: S29129; MUID:93074997
A:Accession: S29130
A:Molecule type: mRNA
A:Residues: 1-384 <TRE>
A:Cross-references: EMBL:X67598
A:Accession: T01068
A:Status: translated from GB/EMBL/DBJ

Query Match 98.4%; Score 313; DB 2; Length 421;
Best Local Similarity 98.3%; Pred. No. 1.9e-29;
Matches 59; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 60
|||||
Db 142 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 201
|||||

RESULT 7
JH0795
calreticulin precursor - California sea hare
N:Alternate names: protein 407
C:Species: Aplysia californica (California sea hare)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0795; B31409; F60977
R:Kennedy, T.E.; Kuhl, D.; Barzilai, A.; Sweatt, J.D.; Kandel, E.R.
Neuron 9, 1013-1024, 1992.
A:Title: Long-term sensitization training in aplysia leads to an increase in calreticulin
A:Reference number: JH0795; MUID:93098937
A:Accession: JH0795
A:Molecule type: mRNA
A:Residues: 1-405 <KEN>
A:Cross-references: GB:S51239; NID:g262053; PIDN:AAB24569.1; PID:g262054
A:Experimental source: abdominal ganglion and antral nervous system
R:Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.; Sweatt, J.D.
Proc. Natl. Acad. Sci. U.S.A. 85, 7008-7012, 1988
A:Title: Sequencing of proteins from two-dimensional gels by using in situ digestion and
tion in Aplysia.
A:Reference number: A94207; MUID:88320566
A:Accession: B31409
A:Molecule type: protein
A:Residues: 'X', 17-28, 'X', 30-31 <KE2>
R:Sweatt, J.D.; Kennedy, T.E.; Wager-Smith, K.; Gawinowicz, M.A.; Barzilai, A.; Karl, K.
Electrophoresis 10, 152-157, 1989
A:Title: Development of a database of amino acid sequences for proteins identified and
A:Reference number: A60977; MUID:89276264
A:Accession: F60977
A:Molecule type: protein
A:Residues: 'X', 17-28, 'X', 30-31 <SWE>
C:Superfamily: calreticulin
C:Keywords: calcium binding; endoplasmic reticulum
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-405/Product: calreticulin #status experimental <MAT>
F:402-405/Region: endoplasmic reticulum retention signal

Query Match 88.4%; Score 281; DB 1; Length 405;
Best Local Similarity 85.0%; Pred. No. 1.2e-25;
Matches 51; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 60
|||||
Db 134 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 193
|||||

RESULT 8
S29130
calreticulin (clone 8) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: S29130; T01068
R:Treves, S.; Zorzato, F.; Pozzan, T.
Biochem. J. 287, 579-581, 1992
A:Title: Identification of calreticulin isoforms in the central nervous system.
A:Reference number: S29129; MUID:93074997
A:Accession: S29130
A:Molecule type: mRNA
A:Residues: 1-384 <TRE>
A:Cross-references: EMBL:X67598
A:Accession: T01068
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-339, <TRW>
A:Cross-references: EMBL:X67598; NID:g64610; PIDN:CAA47867.1; PID:g64611
A:Experimental source: CNS
C:Superfamily: calreticulin
C:Keywords: glycoprotein
F:381-384/Region: endoplasmic reticulum retention signal
F:316/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.4%; Score 278; DB 2; Length 384;
Best Local Similarity 88.3%; Pred. No. 2.5e-25;
Matches 53; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 60
|||||
Db 110 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 169
|||||

RESULT 9
S29129
calreticulin precursor (clone 3) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: S29129
R:Treves, S.; Zorzato, F.; Pozzan, T.
Biochem. J. 287, 579-581, 1992
A:Title: Identification of calreticulin isoforms in the central nervous system.
A:Reference number: S29129; MUID:93074997
A:Accession: S29129
A:Molecule type: mRNA
A:Residues: 1-411 <TRE>
A:Cross-references: EMBL:X67597; NID:g64608; PIDN:CAA47866.1; PID:g64609
C:Superfamily: calreticulin
C:Keywords: glycoprotein
F:1-12/Domain: signal sequence (fragment) #status predicted <SIG>
F:13-411/Product: calreticulin #status predicted <MAT>
F:408-411/Region: endoplasmic reticulum retention signal
F:339/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.4%; Score 278; DB 2; Length 411;
Best Local Similarity 88.3%; Pred. No. 2.7e-25;
Matches 53; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 60
|||||
Db 133 GPGTKKVHVIFNYKGNVLINLKDIRCKDDFTHLYTLVIRPNTYEVKIDNSQVSGSLE 192
|||||

RESULT 10
A56637
calreticulin homolog precursor - fruit fly (Drosophila melanogaster)
N:Alternate names: Ro/SS-A autoantigen/calreticulin homolog
C:Species: Drosophila melanogaster
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 13-Aug-1999
C:Accession: A56637; A37158
R:Smith, M.J.
DNA Seq. 3, 247-250, 1992
A:Title: Nucleotide sequence of a Drosophila melanogaster gene encoding a calreticuli
A:Reference number: A56637; MUID:93208374
A:Accession: A56637
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-406 <SM>
A:Cross-references: GB:X64461; NID:g7685; PIDN:CAA45791.1; PID:g7686
A:Note: Sequence extracted from NCBI backbone (NCBI:128274, NCBI:128275)
R:McCaulliffe, D.P.; Zappi, E.; Lieu, T.S.; Michalak, M.; Sontheimer, R.D.; Capra, J.D.
J. Clin. Invest. 86, 332-335, 1990
A:Title: A human Ro/SS-A autoantigen is the homologue of calreticulin and is highly h
A:Reference number: A37158; MUID:90307981
A:Accession: A37158
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

RESULT 15

calreticulin - beet

calreticulin - beet

C;Species: Beta vulgaris (beet)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text change 20-Jun-2000

C;Accession: T14554

R;Viereck, R.

submitted to the EME

A:Description: Nucle

A: Reference number:

A: Accession: T14554

A: Status: preliminary

A: Molecule type: mpn

A: MOLECULE TYPE: IRR
A: Residues: 1-416 <V

A; residues: 1-416 < A: Cross-references:

A: Experimental sources:

A; Experimental source
C: Superfamily: cal ro

C;superfamily: calce

C; Keywords: calcium

Query Match

Best Local Similarity

Matches 33; Conservation

МАШИНЫ, КОМПЬЮТЕРЫ

1 GPGTKKHVTFNY

Y T GPGTKKVVHVFNFY

144 GYSTKKVHATFNV

L44 GYSTKRVHAIFFNYI

Search completed: October 21, 2001, 03:07:45

Job time: 340 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:08:25 ; Search time 28.98 Seconds
(without alignments)
70.922 Million cell updates/sec

Title: US-09-807-148-8
Perfect score: 318
Sequence: 1 GPGTKVHVIFNYKGNVLI.....PDNTYEVKIDNSQVESGSL6 60

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	318	100.0	416	1	P14211 mus musculus
2	318	100.0	416	1	CRTC_MOUSE
3	318	100.0	417	1	CRTC_RAT
4	318	100.0	418	1	CRTC_HUMAN
5	313	98.4	400	1	CRTC_RABIT
6	313	98.4	421	1	CRT1_BOVIN
7	269	84.6	405	1	CRT2_BOVIN
8	242	76.1	388	1	CRTC_DROME
9	227	71.4	393	1	RAL1_ONCOCERCA
10	210.5	66.2	395	1	CRTC_SCHMA
11	199	62.6	420	1	CRTC_CAEEL
12	184	57.9	416	1	CRTC_CHLRE
13	179	56.3	421	1	CRTC_BETVU
14	177.5	55.8	424	1	CRTC_PRUAR
15	177	55.7	415	1	CRTC_DICDI
16	177	55.7	424	1	CRTC_RICCO
17	172	54.1	425	1	CRTC_ARATH
18	171	53.8	424	1	CRTC_ORYSA
19	170	53.5	420	1	CRTC_MAIZE
20	169	53.1	416	1	CRTC_NICPL
21	166	52.2	401	1	CRTC_EUGGR
22	166	52.2	424	1	CRTC_ATH
23	154	48.4	416	1	CRTC_BERST
24	117	36.8	592	1	CALX_HUMAN
25	116	36.5	591	1	CALX_MOUSE
26	116	36.5	591	1	CALX_RAT
27	115	36.2	593	1	CALX_CANFA
28	107	33.6	610	1	CALG_HUMAN
29	106	33.3	611	1	CALG_MOUSE
30	98	30.8	619	1	CALX_CAEEL
31	97.5	30.7	560	1	CALX_SCHPO
32	87	27.4	546	1	CALX_SOYBN
33	82	25.8	530	1	CAX1_ARATH

34 77 24.2 540 1 CALX_HELTU
35 551 1 CALX_PEA
36 74 23.3 528 1 CAX2_ARATH
37 64.5 20.3 540 1 MTAL_ACICA
38 61.5 19.3 582 1 SYN2_HUMAN
39 61 19.2 474 1 LAM3_MOUSE
40 61 19.2 592 1 LAM2_MOUSE
41 60.5 19.0 586 1 SYN2_RAT
42 59.5 18.7 220 1 PCD9_YEAST
43 58.5 18.4 362 1 YCY5_YEAST
44 57.5 18.1 126 1 RR13_PORPU
45 57 17.9 339 1 REP_BACAM

ALIGNMENTS

RESULT_1

ID CRTC_MOUSE STANDARD; PRT; 416 AA.
AC P14211;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60).
GN CALR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-48 AND 129-161.
RX STRAIN=BALB/C; TISSUE=Liver;
RC MEDLINE=90059955; PubMed=2583110;
RA Smith M.J., Koch G.L.E.;
RT "Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein.";
RL EMBO J. 8:3581-3586(1989).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93013037; PubMed=1398135;
RA Mazzarella R.A., Gold P., Cunningham M., Green M.;
RT "Determination of the sequence of an expressible cDNA clone encoding Erp60/calregulin by the use of a novel nested set method.";
RL Gene 120:217-225(1992).
[3]
RN SEQUENCE OF 18-38.
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14926; CAA33053.1; -;
DR EMBL; M92988; AAA37569.1; -;
DR PIR; S06763; S06763.
DR PIR; JC1444; JC1444.
DR SWISS-2DPAGE; P14211; MOUSE.
DR MGI; MGI:88252; Calr.

```

DR InterPro: IPR000886;
DR Pfam: PF001580;
DR PRINTS: PF00262; calreticulin; 1.
DR PROSITE: PS00626; CALRETICULIN.
DR PROSITE: PS00014; ER_TARGET; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 416 CALRETICULIN.
FT DOMAIN 18 197 N-DOMAIN.
FT DOMAIN 198 308 P-DOMAIN.
FT DOMAIN 309 416 C-DOMAIN.
FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT REPEAT 191 202 1-1.
FT REPEAT 210 221 1-2.
FT REPEAT 227 238 1-3.
FT REPEAT 244 255 1-4.
FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT REPEAT 259 269 2-1.
FT REPEAT 273 283 2-2.
FT REPEAT 287 297 2-3.
FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
FT DISULFID 137 163 BY SIMILARITY.
FT SITE 413 416 PREVENT SECRETION FROM ER.
SQ SEQUENCE 416 AA: 47994 MW: 24C03B00913408D8 CRC64;

Query Match 100.0%; Score 318; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 3.4e-31;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGTKVHVIFNYKGNVLINKDIRCKDDEFTLYLIVRPDNTYEVKIDNSQVSGSLE 60
DB 138 GPGTKVHVIFNYKGNVLINKDIRCKDDEFTLYLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 2
ID CRTCLAT STANDARD; PRT: 416 AA.
AC P18(18; P10452;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (CALBP)
DE (CALCIUM-BINDING PROTEIN 3) (CABP3).
GN CALR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain cortex;
RX MEDLINE=90370496; PubMed=2395661;
RA Murthy K.K., Banville D., Srikanth C.B., Carrier F., Bell A.,
RA Holmes C., Patel Y.C.;
RT "Structural homology between the rat calreticulin gene product and
the Onchocerca volvulus antigen Ral-1."
RL Nucleic Acids Res. 18:4933-4933(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE=93202172; PubMed=8453984;
RA Nakamura M., Moriya M., Baba T., Michikawa Y., Yamanobe T., Arai K.,
RA Okinaga S., Kobayashi T.;
RT "An endoplasmic reticulum protein, calreticulin, is transported into
the acrosome of rat sperm."
RL Exp. Cell Res. 205:101-110(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Liver;

```

```

RX MEDLINE=95181573; PubMed=7876339;
RA Soennichsen B., Fuehlekug J., van Nguyen P., Diekmann W.,
RA Robinson D.G., Mieskes G.;
RT "Retention and retrieval: both mechanisms cooperate to maintain
calreticulin in the endoplasmic reticulum."
RL J. Cell Sci. 107:2705-2717(1994).
RN [4]
RP SEQUENCE OF 270-358 FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA Lone Y.C., Bailly A., Latruffe N.;
RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 18-29.
RX MEDLINE=91054414; PubMed=2241926;
RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,
RA MacLennan D.H., Meldolesi J., Pozzan T.;
RT "Calreticulin is a candidate for a calsequestrin-like function in
Ca2(+)-storage compartments (calciosomes) of liver and brain."
RL Biochem. J. 271:473-480(1990).
RN [6]
RP SEQUENCE OF 18-32.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Testis;
RX MEDLINE=92360010; PubMed=1497655;
RA Nakamura M., Michikawa Y., Baba T., Okinaga S., Arai K.;
RT "Calreticulin is present in the acrosome of spermatids of rat
testis."
RL Biochem. Biophys. Res. Commun. 186:668-673(1992).
RN [7]
RP SEQUENCE OF 18-32.
RC STRAIN-LEC; TISSUE=Liver;
RX MEDLINE=94072621; PubMed=8251535;
RA Yokoi T., Nagayama S., Kajiwarra R., Kawaguchi Y., Horiuchi R.,
RA Kamataki T.;
RT "Identification of protein disulfide isomerase and calreticulin as
autoimmune antigens in LEC strain of rats."
RL Biochim. Biophys. Acta 1158:339-344(1993).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC -!- LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -!- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE D-BETA-
CC HYDROXYBUTYRATE DEHYDROGENASE.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: D78308; BAA11345.1;
EMBL: X53363; CAA37446.1;
EMBL: X13702; CAA31987.1; ALT_SEQ.
EMBL: X79327; CAA55890.1;
PIR: S04867; S04867.
PIR: S11205; S11205.
PIR: S13045; S13045.
PIR: A49176; A49176.
PIR: S45036; S45036.
PIR: JH0819; JH0819.
PIR: JH0819; JH0819.
InterPro: IPR001580;
Pfam: PF00262; calreticulin; 1.
PRINTS: PR00626; CALRETICULIN.
PROSITE: PS00014; ER_TARGET; 1.
PROSITE: PS00803; CALRETICULIN_1; 1.
PROSITE: PS00804; CALRETICULIN_2; 1.
PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 416 CALRETICULIN.

```

FT DOMAIN 18 197 N-DOMAIN.
 FT DOMAIN 198 308 P-DOMAIN.
 FT DOMAIN 309 416 C-DOMAIN.
 FT DOMAIN 191 255 4 X APPROXIMATE REPEATS.
 FT REPEAT 191 202 1-1.
 FT REPEAT 210 221 1-2.
 FT REPEAT 227 238 1-3.
 FT REPEAT 244 255 1-4.
 FT DOMAIN 259 297 3 X APPROXIMATE REPEATS.
 FT REPEAT 259 269 2-1.
 FT REPEAT 273 283 2-2.
 FT REPEAT 287 297 2-3.
 FT DOMAIN 351 407 ASP/GLU/LYS-RICH.
 FT DISULFID 137 163 BY SIMILARITY.
 FT SITE 413 416 PREVENT SECRETION FROM ER.
 SQ SEQUENCE 416 AA; 47995 MW; 266713CED31A2970 CRC64;

Query Match 100.0%; Score 318; DB 1; Length 416;
 Best Local Similarity 100.0%; Pred. No. 3.4e-31;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGTKVHVIFNYGKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 |||||
 Db 138 GPGTKVHVIFNYGKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 197

RESULT 3
 CRTC_HUMAN
 ID CRTC_HUMAN STANDARD; PRT; 417 AA.
 AC P27797;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (52 KDA
 DE RIBONUCLEOPROTEIN AUTOANTIGEN RO/SS-A).
 GN CALR OR CRTC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92013129; PubMed=1919005;
 RA Rokeach L.A., Haselby J.A., Meilof J.F., Smeenk R.J., Unnasch T.R.,
 RA Greene B.M., Hoch S.O.;
 RT "Characterization of the autoantigen calreticulin.";
 RL J. Immunol. 147:3031-3039(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90237213; PubMed=2332496;
 RA McCauliffe D.P., Lux F.A., Lieu T.S., Sanz I., Hanke J., Newkirk M.M.,
 RA Bachinski L.L., Itoh Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,
 RA Capra J.D.;
 RT "Molecular cloning, expression, and chromosome 19 localization of a
 human Ro/SS-A autoantigen.";
 RL J. Clin. Invest. 85:1379-1391(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92129342; PubMed=1733953;
 RA McCauliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;
 RT "The 5'-flanking region of the human calreticulin gene shares
 RT homology with the human GRP78, GRP94, and protein disulfide isomerase
 RT promoters.";
 RL J. Biol. Chem. 267:2557-2562(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J., McCreedy P., Stilwagen S., Ramirez M., Carrano A.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 18-36.
 RX MEDLINE=92002034; PubMed=1911778;
 RA Rojiani M.V., Finlay B.B., Gray V., Dedhar S.;

RT "In vitro interaction of a polypeptide homologous to human Ro/SS-A
 RT antigen (calreticulin) with a highly conserved amino acid sequence in
 RT the cytoplasmic domain of integrin alpha subunits.";
 RL Biochemistry 30:9859-9866(1991).
 RN [6]
 RP SEQUENCE OF 18-32.
 RX MEDLINE=90380058; PubMed=2400400;
 RA Krause K.H., Simmerman H.K.B., Jones L.R., Campbell K.P.;
 RT "Sequence similarity of calreticulin with a Ca2(+)-binding protein
 RT that co-purifies with an Ins(1,4,5)P3-sensitive Ca2+ store in HL-60
 RT cells.";
 RL Biochem. J. 270:545-548(1990).
 RN [7]
 RP SEQUENCE OF 18-28.
 RC TISSUE=Liver;
 RX MEDLINE=93162045; PubMed=1286669;
 RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
 RA Appel R.D., Hughes G.J.;
 RT "Human liver protein map: a reference database established by
 RT microsequencing and gel comparison.";
 RL Electrophoresis 13:992-1001(1992).
 RN [8]
 RP PARTIAL SEQUENCE OF 25-34; 56-62; 208-221 AND 273-278.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [9]
 RP SEQUENCE OF 18-26.
 RC TISSUE=Colon carcinoma;
 RX MEDLINE=97295306; PubMed=9150948;
 RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
 RT "A two-dimensional gel database of human colon carcinoma proteins.";
 RL Electrophoresis 18:605-613(1997).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE THE RO AUTOANTIGEN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M84739; AAA51916.1; -
 DR EMBL; M32294; AAA36582.1; -
 DR EMBL; AD000092; AAB5176.1; -
 DR PIR; A37047; A37047.
 DR PIR; S11475; S11475.
 DR PIR; A42330; A42330.
 DR PIR; A46452; A46452.
 DR SWISS-2DPAGE; P27797; HUMAN.
 DR Aarhuus/Ghent-2DPAGE; 9401; IEF.
 DR HSC-2DPAGE; P27797; HUMAN.
 DR MIM; 109091; -
 DR InterPro; IPR000886; -
 DR InterPro; IPR001580; -
 DR Pfam; PF00262; Calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.

FT SIGNAL 1 17
 FT CHAIN 18 417
 FT DOMAIN 18 197
 FT DOMAIN 198 308
 FT DOMAIN 309 417
 FT DOMAIN 191 255
 FT REPEAT 191 202
 FT REPEAT 210 221
 FT REPEAT 227 238
 FT REPEAT 244 255
 FT DOMAIN 259 297
 FT REPEAT 259 269
 FT REPEAT 273 283
 FT REPEAT 287 297
 FT DOMAIN 351 408
 FT DISULFID 137 163
 FT SITE 414 417
 FT CONFLICT 35 35
 FT SEQUENCE 417 AA; 48141 MW; BC37C3C0F1054FB2 CRC64;
 SQ

Query Match 100.0%; Score 318; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 3.4e-31;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNYKGNVLINCKDDEFFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
 DB 138 GPCTKKVHVIFNYKGNVLINCKDDEFFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 4
 CRTC-RABIT
 ID CRTC-RABIT STANDARD; PRT; 418 AA.
 AC P1253;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60).
 GN CALR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Slow-twitch skeletal muscle;
 RX MEDLINE=90094320; PubMed=2600080;
 RA Fliegel L., Burns K., MacLennan D.H., Reithmeier R.A.F., Michalak M.;
 RT "Molecular cloning of the high affinity calcium-binding protein
 (calreticulin) of skeletal muscle sarcoplasmic reticulum";
 RL J. Biol. Chem. 264:21522-21528(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fast-twitch skeletal muscle;
 RX MEDLINE=91282795; PubMed=2059224;
 RA Fliegel L., Michalak M.;
 RT "Fast-twitch and slow-twitch skeletal muscles express the same
 isoform of calreticulin";
 RL Biochem. Biophys. Res. Commun. 177:979-984(1991).
 RN [3]
 RP SEQUENCE OF 18-36.
 RX MEDLINE=91054414; PubMed=2241926;
 RA Treves S., de Mattei M., Lanfretti M., Villa A., Green N.M.,
 MacLennan D.H., Meldolesi J., Pozzan T.;
 RT "Calreticulin is a candidate for a calsequestrin-like function in
 Ca2(+)-storage compartments (calciosomes) of liver and brain";
 RL Biochem. J. 271:473-480(1990).
 RN [4]
 RP SEQUENCE OF 18-46.
 RX MEDLINE=91201375; PubMed=2016321;
 RA Milner R.E., Baksh S., Shemanko C., Carpenter M.R., Smillie L.,
 Vance J.E., Opas M., Michalak M.;
 RT "Calreticulin, and not calsequestrin, is the major calcium binding

RT protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic
 reticulum";
 RL J. Biol. Chem. 266:7155-7165(1991).
 RN [5]
 RP PARTIAL SEQUENCE.
 RC TISSUE=Lung;
 RX MEDLINE=92002038; PubMed=1911780;
 RA Guan S., Falick A.M., Williams D.E., Cashman J.R.;
 RT "Evidence for complex formation between rabbit lung flavin-containing
 monooxygenase and calreticulin";
 RL Biochemistry 30:9892-9900(1991).
 CC -|- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC -|- LOW AFFINITY CALCIUM-BINDING SITES.
 CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -|- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J05138; AAA31188.1;
 DR PIR: A34154; A34154.
 DR PIR: C33208; C33208.
 DR PIR: D33208; D33208.
 DR PIR: E33208; E33208.
 DR PIR: F33208; F33208.
 DR PIR: S13046; S13046.
 DR PIR: S13047; S13047.
 DR InterPro: IPR000886;
 DR InterPro: IPR001580;
 DR Pfam: PF00262; calreticulin; 1.
 DR PRINTS: PR00626; CALRETICULIN.
 DR PROSITE: PS00014; ER_TARGET; 1.
 DR PROSITE: PS00803; CALRETICULIN_1; 1.
 DR PROSITE: PS00804; CALRETICULIN_2; 1.
 DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 418
 FT DOMAIN 18 197
 FT DOMAIN 198 308
 FT DOMAIN 309 418
 FT DOMAIN 191 255
 FT REPEAT 191 202
 FT REPEAT 210 221
 FT REPEAT 227 238
 FT REPEAT 244 255
 FT DOMAIN 259 297
 FT REPEAT 259 269
 FT REPEAT 273 283
 FT REPEAT 287 297
 FT DOMAIN 351 408
 FT DISULFID 137 163
 FT SITE 415 418
 FT VARIANT 35 35
 FT CONFLICT 90 90
 FT SEQUENCE 418 AA; 48275 MW; B6082B689DC763A6 CRC64;
 SQ

Query Match 100.0%; Score 318; DB 1; Length 418;
 Best Local Similarity 100.0%; Pred. No. 3.4e-31;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNYKGNVLINCKDDEFFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
 DB 138 GPCTKKVHVIFNYKGNVLINCKDDEFFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 5

OC	CRTI_BOVIN	STANDARD;	PRT;	400 AA.
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=93385184; PubMed=8373827;			
RA	Liu N., Fine R.E., Johnson R.J.;			
RT	*Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin.*;			
RL	Biochim. Biophys. Acta 1202:70-76(1993).			
CC	-!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND LOW AFFINITY CALCIUM-BINDING SITES.			
CC	-!- SUBUNIT: MONOMER (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.			
CC	-!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
DR	EMBL; L13462; AAC37307.1;			
DR	InterPro; IPR000886;			
DR	InterPro; IPR001580;			
DR	Pfam; PF00262; calreticulin_1.			
DR	PRINTS; PR00626; CALRETICULIN.			
DR	PROSITE; PS00014; ER_TARGET; 1.			
DR	PROSITE; PS00803; CALRETICULIN_1; 1.			
DR	PROSITE; PS00804; CALRETICULIN_2; 1.			
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.			
KW	Endoplasmic reticulum; Calcium-binding; Repeat; Signal.			
FT	SIGNAL 1 34 POTENTIAL.			
FT	CHAIN 35 421 CALRETICULIN, BRAIN ISOFORM 2.			
FT	DOMAIN 35 201 N-DOMAIN.			
FT	DOMAIN 202 312 P-DOMAIN.			
FT	DOMAIN 313 421 C-DOMAIN.			
FT	DOMAIN 195 259 4 X APPROXIMATE REPEATS.			
FT	REPEAT 195 206 1-1.			
FT	REPEAT 214 225 1-2.			
FT	REPEAT 231 242 1-3.			
FT	REPEAT 248 259 1-4.			
FT	DOMAIN 263 301 3 X APPROXIMATE REPEATS.			
FT	REPEAT 263 273 2-1.			
FT	REPEAT 277 287 2-2.			
FT	REPEAT 291 301 2-3.			
FT	DOMAIN 366 411 ASP/GLU/LYS-RICH.			
FT	DISULFID 141 167 BY SIMILARITY.			
FT	CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	SITE 418 421 PREVENT SECRETION FROM ER.			
SQ	SEQUENCE 421 AA; 48812 MW; 0257E959F71528BC CRC64;			

Query Match 98.4%; Score 313; DB 1; Length 421;
Best Local Similarity 98.3%; Pred. No. 1.4e-30;
Matches 59; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	GPCTKKVHVFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEYKIDNSQVESGSLE 60
Db	142	GPCTKKVHVFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEYKIDNSQVESGSLE 201

RESULT 6

ID	CRT2_BOVIN	STANDARD;	PRT;	421 AA.
ID	F42918;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	CALRETICULIN, BRAIN ISOFORM 2 PRECURSOR (CRP55) (CALREGULIN) (HACBP).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OX	NCBI_TaxID=9913;			
RN	[1]			
RN	SEQUENCE.			
RC	TISSUE=Brain;			
RX	MEDLINE=94183174; PubMed=8135753;			
RA	Matsuoka K., Seta K., Yamakawa Y., Okuyama T., Shinoda T., Isobe T.;			
RT	"Covalent structure of bovine brain calreticulin.";			
RL	Biochem. J. 298:435-442(1994).			
CC	-!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND LOW AFFINITY CALCIUM-BINDING SITES.			
CC	-!- SUBUNIT: MONOMER.			
CC	-!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.			
CC	-!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.			
DR	InterPro; IPR000886;			
DR	InterPro; IPR001580;			
DR	Pfam; PF00262; calreticulin_1.			
DR	PRINTS; PR00626; CALRETICULIN.			
DR	PROSITE; PS00014; ER_TARGET; 1.			
DR	PROSITE; PS00803; CALRETICULIN_1; 1.			
DR	PROSITE; PS00804; CALRETICULIN_2; 1.			
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.			
KW	Endoplasmic reticulum; Calcium-binding; Repeat; Glycoprotein.			
FT	DOMAIN 1 18			

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93208374; PubMed=1296819;
 RX Smith M.J.;
 RA "Nucleotide sequence of a Drosophila melanogaster gene encoding a
 RT calreticulin homologue";
 RL DNA Seq. 3:247-250(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei I., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 91-124 AND 182-220.
 RX MEDLINE=90307981; PubMed=2365822;
 RA McCalliffe D.P., Zappi E., Lieu T.S., Michalak M., Sontheimer R.D.,
 RA Capra J.D.;
 RT "A human Ro/SS-A autoantigen is the homologue of calreticulin and is
 RT highly homologous with onchocercal RAL-1 antigen and an aplasia
 RT 'memory molecule'";
 RL J. Clin. Invest. 86:332-335(1990).
 CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X64461; CAA45791.1; -;
 DR EMBL; AF003683; AAF54416.1; -;
 DR EIR; A37158; A37158;
 DR FlyBase; FBgn0005585; Crc.
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 406 CALRETICULIN.
 FT CONFLICT 107 107 G -> A (IN REF. 3).
 FT CONFLICT 184 184 V -> L (IN REF. 3).
 SQ SEQUENCE 406 AA; 46808 MW; 65D72C69D0BEC427 CRC64;
 Query: Match 84.6%; Score 269; DB 1; Length 406;
 Best Local Similarity 83.3%; Pred. No. 2.8e-25;
 Matches 50; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 GPOTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYTLIVRPDNTYEVKIDNSOVESGSL 60
 DB 138 GPOTKKVHVIFSYKGNHLISKDIRCKDDVYTHFTLYTLIVRPDNTYEVLDNEKVESGNLE 197
 RESULT 8
 RALL ONCVO STANDARD; PRT; 388 AA.
 ID RALL ONCVO STANDARD; PRT; 388 AA.
 AC P11012;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE RAL-1 PROTEIN PRECURSOR (41 KDA-LARVAL ANTIGEN).
 GN RAL1.
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94341871; PubMed=7520419;
 RA Rokeach L.A., Zimmerman P.A., Unnasch T.R.;
 RT "Epitopes of the Onchocerca volvulus RAL1 antigen, a member of the
 RT calreticulin family of proteins, recognized by sera from patients
 RT with onchocerciasis";
 RL Infect. Immun. 62:3696-3704(1994).
 RN [2]
 RP SEQUENCE OF 53-388 FROM N.A.
 RX MEDLINE=88273584; PubMed=2455736;
 RA Unnasch T.R., Gallin M.Y., Soboslay P.T., Erttmann K.D., Greene B.M.;
 RT "Isolation and characterization of expression cDNA clones encoding
 RT antigens of Onchocerca volvulus infective larvae";
 RL J. Clin. Invest. 82:262-269(1988).
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M20565; AAA59056.1; -;
 DR EIR; A32507; A32507; -;
 DR InterPro; IPR001580; -;

DR PFam: PF00262; calreticulin; 1.
 DR PRINTS: PR00626; CALRETICULIN
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Calcium-binding; Repeat; Antigen; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 388 RAL-1 PROTEIN.
 FT DOMAIN 189 253 4 X APPROXIMATE REPEATS.
 FT REPEAT 189 200 1-1.
 FT REPEAT 208 219 1-2.
 FT REPEAT 225 236 1-3.
 FT REPEAT 242 253 1-4.
 FT DOMAIN 257 295 3 X APPROXIMATE REPEATS.
 FT REPEAT 257 267 2-1.
 FT REPEAT 271 281 2-2.
 FT REPEAT 285 295 2-3.
 FT DOMAIN 335 388 ARG/LYS-RICH (BASIC).
 FT DISULFID 135 161 BY SIMILARITY.
 FT SEQUENCE 388 AA; 45298 MW; 9537F298A2D31CD6 CRC64;
 Query Match 76.1%; Score 242; DB 1; Length 388;
 Best Local Similarity 75.0%; Pred. No. 5e-22;
 Matches 45; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
 QY 1 GPGTKKVVHVIENYKGNVLINKDKRDETHLYTLVIRPDNTYEVKIDNSQVESGSL 60
 DB 136 GPGTKKVVHVIENYKGNVLINKDKRDETHLYTLVIRPDNTYEVKIDNSQVESGSL 60
 RESULT 9
 ID CRTCSCHMA STANDARD; PRT; 393 AA.
 AC Q06814; Q26562;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CALRETICULIN PRECURSOR (SM4 PROTEIN).
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigoida; Schistosomatoidea; Schistosomatidae;
 OC Schistosoma.
 OX NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PUERTO RICAN;
 RX MEDLINE=93165070; PubMed=8433712;
 RA Khalife J., Trottein F., Schacht A.-M., Godin C., Pierce R.J.,
 RA Capron A.;
 RT "Cloning of the gene encoding a Schistosoma mansoni antigen
 homologous to human RO/SS-A autoantigen.";
 RL Mol. Biochem. Parasitol. 57:193-202(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PUERTO RICAN;
 RX MEDLINE=94187805; PubMed=8139623;
 RA Khalife J., Pierce R.J., Godin C., Capron A.;
 RT "Cloning and sequencing of the gene encoding Schistosoma mansoni
 calreticulin.";
 RL Mol. Biochem. Parasitol. 62:313-315(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=92329978; PubMed=1627827;
 RA Smith M.J.;
 RT "A C. elegans gene encodes a protein homologous to mammalian
 calreticulin.";
 RL DNA Seq. 2:235-240(1992).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M93097; AAA29854.1; -
 DR EMBL; L24159; AAA19024.1; -
 DR HSP; P00268; 4RXN.
 DR InterPro; IPR000886; -
 DR InterPro; IPR001580; -
 DR PRINTS: PR00626; calreticulin; 1.
 DR PROSITE; PS00626; CALRETICULIN.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 393 CALRETICULIN.
 FT DOMAIN 189 254 4 X 12 AA APPROXIMATE REPEATS.
 FT REPEAT 189 200 1-1.
 FT REPEAT 209 220 1-2.
 FT REPEAT 225 236 1-3.
 FT REPEAT 243 254 1-4.
 FT DOMAIN 257 295 3 X 11 AA APPROXIMATE REPEATS.
 FT REPEAT 257 267 2-1.
 FT REPEAT 271 281 2-2.
 FT REPEAT 285 295 2-3.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 135 161 BY SIMILARITY.
 FT SITE 390 393 PREVENT SECRETION FROM ER.
 FT CONFLICT 89 90 MV -> IL (IN REF. 2).
 FT CONFLICT 188 207 MISSING (IN REF. 2).
 FT CONFLICT 378 378 Y -> D (IN REF. 2).
 FT SEQUENCE 393 AA; 45397 MW; 45F59857C21940D2 CRC64;
 Query Match 71.4%; Score 227; DB 1; Length 393;
 Best Local Similarity 73.3%; Pred. No. 3.3e-20;
 Matches 44; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
 QY 1 GPGTKKVVHVIENYKGNVLINKDKRDETHLYTLVIRPDNTYEVKIDNSQVESGSL 60
 DB 136 GPGTKKVVHVIENYKGNVLINKDKRDETHLYTLVIRPDNTYEVKIDNSQVESGSL 60
 RESULT 10
 ID CRTCSAEEL STANDARD; PRT; 395 AA.
 AC P27798;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE CALRETICULIN PRECURSOR.
 GN CRT-1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=92329978; PubMed=1627827;
 RA Smith M.J.;
 RT "A C. elegans gene encodes a protein homologous to mammalian
 calreticulin.";
 RL DNA Seq. 2:235-240(1992).
 CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

2

Search completed: October 21, 2001, 03:08:25
Job time: 275 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	294	92.5	318	13	Q9ptx7	lampetra re
2	281	88.4	405	5	Q26288	aplysia cal
3	278	87.4	343	13	Q91711	xenopus lae
4	278	87.4	411	13	Q91710	xenopus lae
5	270	84.9	421	5	Q9U6S0	strongyloce
6	269	84.6	406	5	Q9U16	drosohilla
7	267	84.0	5	5	Q16893	amblyomma a
8	262	82.4	410	5	Q16893	amblyomma a
9	259	81.4	419	13	Q98984	rana rugosa
10	249	78.3	387	13	Q9PUC1	brachydanio
11	243	76.4	375	5	Q97372	dirofilaria
12	240.5	75.6	403	5	Q18478	litomosoides
13	237	74.5	350	5	Q76961	nectator ame
14	237	74.5	350	5	Q26514	schistosoma
15	217	68.2	321	13	Q45034	schistosoma
16	179	56.3	421	10	Q9U5G0	epratretus
17	179	55.7	321	10	Q43712	zea mays (m
18	176	55.3	412	10	Q41799	zea mays (m
19	176	55.3	415	10	Q40040	hordeum vul

RESULT 2

1. What is the purpose of the study?
2. What are the research objectives?
3. What is the research methodology?
4. What are the results of the study?
5. What are the conclusions of the study?
6. What are the limitations of the study?
7. What are the implications of the study?
8. What are the future research directions?
9. What are the contributions of the study?
10. What are the key findings of the study?

DR DR

Query Match 84.0%; Score 267; DB 5; Length 410;
Best Local Similarity 83.3%; Pred. No. 6.6e-23;

```
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PS00626; CALRETICULIN.
DR ProDom: PD001866; -; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00805; ER_TARGET; UNKNOWN_1.
DR PROSITE: PS0014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 419 AA; 48658 MW; 2C857036769673BF CRC64;

Query Match 82.4%; Score 262; DB 13; Length 419;
Best Local Similarity 83.3%; Pred. No. 2.5e-22;
Matches 50; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 139 GPPTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGNLE 198

RESULT 9
Q9PUC1 PRELIMINARY; PRT; 417 AA.
AC Q9PUC1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubinstein A.L., Lee D., Luo R., Henion P.D., Halpern M.E.;
RT "Genes dependent on zebrafish cyclops function identified by AFLP
RT Differential Gene Expression Screen.";
RL Genesis 0:0-(1999).
DR EMBL: AF195882; AAF13700.1; -
DR InterPro: IPR000886; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PS00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS0014; ER_TARGET; UNKNOWN_1.
DR PROSITE: PS0014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 417 AA; 48723 MW; 2000C5B4004699B6 CRC64;

Query Match 81.4%; Score 259; DB 13; Length 417;
Best Local Similarity 81.7%; Pred. No. 5.6e-22;
Matches 49; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 GYSTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNEKVESGSL 197

RESULT 10
Q97372 PRELIMINARY; PRT; 387 AA.
AC Q97372;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Dirofilaria immitis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Dirofilaria.
OX NCBI_TaxID=6287;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=99094497; PubMed=9879888;
RA Tsuji N., Morales T.H., Ozols V.V., Carmody A.B., Chandrashekar R.;
RT "Molecular characterization of a calcium-binding protein from the
RT filarial parasite Dirofilaria immitis.";
RL Mol. Biochem. Parasitol. 97:69-79(1998).
DR EMBL: AF052978; AAD03405.1; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PS00626; CALRETICULIN.
DR ProDom: PD001866; -; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 387 CALRETICULIN.
SQ SEQUENCE 387 AA; 44941 MW; E7741BF6AAFA5885 CRC64;

Query Match 78.3%; Score 249; DB 5; Length 387;
Best Local Similarity 76.7%; Pred. No. 7.3e-21;
Matches 46; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 GPGTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 195

RESULT 11
O18478 PRELIMINARY; PRT; 375 AA.
AC O18478;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE RAL-1 PROTEIN (FRAGMENT).
OS Litomosoides sigmodontis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Litomosoides.
OX NCBI_TaxID=42156;
RN [1]
RP SEQUENCE FROM N.A.
RA MacLennan K., Hoffman W.H., Taylor D.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ001621; CAA04877.1; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PS00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
FT NON_TER 375 375
SQ SEQUENCE 375 AA; 43842 MW; 03F7642FBFF7A5B8 CRC64;

Query Match 76.4%; Score 243; DB 5; Length 375;
Best Local Similarity 75.0%; Pred. No. 3.5e-20;
Matches 45; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 GPGTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 GPGTKKVHVIFNYKGNVLINCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 195

RESULT 12
O76961 PRELIMINARY; PRT; 403 AA.
AC O76961;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN PRECURSOR.
```



```
GN CRT.
OS Necator americanus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoidae; Ancylostomatidae; Bunostominae; Necator.
OX NCBI_TaxID=51031;
[1]
RN SEQUENCE FROM N.A.
RA Pritchard D.I., Brown A., McElroy P., Loukas A., Girdwood K.,
RA Berry C., Fulkrug R., Beck E.;
RT "Calreticulin is a hookworm allergen.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ006790; CAA07254.1; -.
DR InterPro: IPR000886; -.
DR InterPro: IPR001580; -.
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PRODOM: PD001866; -.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW Signal; Allergen.
FT SIGNAL 1 16 POTENTIAL.
SQ SEQUENCE 403 AA; 46833 MW; 21F3B80515487B6F CRC64;

Query Match 75.6%; Score 240.5; DB 5; Length 403;
Best Local Similarity 78.3%; Pred. No. 7.3e-20;
Matches 47; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 1 GPGTKKHVHVIENYKGNVINKDKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
Db 135 GP-PKKVHDIIFYSGKGNHLIKKIDCKDDKTHLYTLIVRPDNTYEVKIDNSQVSGSLE 193

RESULT 13
Q26514 PRELIMINARY; PRT; 350 AA.
AC Q26514;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CALRETICULIN (FRAGMENT).
GN RAL-1.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
OC Trematoda; Digenea; Strigeida; Schistosomatoidae; Schistosomatidae;
OX Schistosoma.
OX NCBI_TaxID=6182;
[1]
RN SEQUENCE FROM N.A.
RA Huggins M.C., Moloney N.A.;
RC STRAIN=CHINESE;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: M80524; AAA2917.1; -.
DR InterPro: IPR000886; -.
DR InterPro: IPR001580; -.
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 1.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 1
SQ SEQUENCE 350 AA; 40385 MW; 30FAB4E8BB685D1C CRC64;

Query Match 74.5%; Score 237; DB 5; Length 350;
Best Local Similarity 73.3%; Pred. No. 1.6e-19;
Matches 44; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 GPGTKKHVHVIENYKGNVINKDKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
Db 135 GP-PKKVHDIIFYSGKGNHLIKKIDCKDDKTHLYTLIVRPDNTYEVKIDNSQVSGSLE 193

RESULT 15
Q9U5G0 PRELIMINARY; PRT; 321 AA.
AC Q9U5G0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Eptatretus burgeri (Inshore hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptatretninae; Eptatretus.
OX NCBI_TaxID=7764;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20063780; PubMed=10594174;
RA Kuraku S., Hoshiyama D., Kato K., Suga H., Miyata T.;
RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded
genes.";
RL J. Mol. Evol. 49:729-735(1999).
DR EMBL: AB025323; BAA88476.1; -.
DR InterPro: IPR000886; -.
DR InterPro: IPR001580; -.
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PRODOM: PD001866; -.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 1
```

```
Db 91 GMATKRIHVIENYKGNHLIKKIDCKDDKTHLYTLIVRPDNTYEVKIDNSQVSGSLE 150

RESULT 14
O45034 PRELIMINARY; PRT; 396 AA.
AC O45034;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
OC Trematoda; Digenea; Strigeida; Schistosomatoidae; Schistosomatidae;
OX Schistosoma.
OX NCBI_TaxID=6182;
[1]
RN SEQUENCE FROM N.A.
RA STRAIN=PHILIPPINE (MINDORO);
RA Scott J.C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF044408; AAC00515.1; -.
DR InterPro: IPR000886; -.
DR InterPro: IPR001580; -.
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PRODOM: PD001866; -.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 396 AA; 45813 MW; C57394C6FB4CD77B CRC64;

Query Match 74.5%; Score 237; DB 5; Length 396;
Best Local Similarity 73.3%; Pred. No. 1.8e-19;
Matches 44; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 GPGTKKHVHVIENYKGNVINKDKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
Db 137 GMATKRIHVIENYKGNHLIKKIDCKDDKTHLYTLIVRPDNTYEVKIDNSQVSGSLE 196

RESULT 15
Q9U5G0 PRELIMINARY; PRT; 321 AA.
AC Q9U5G0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Eptatretus burgeri (Inshore hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptatretninae; Eptatretus.
OX NCBI_TaxID=7764;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20063780; PubMed=10594174;
RA Kuraku S., Hoshiyama D., Kato K., Suga H., Miyata T.;
RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded
genes.";
RL J. Mol. Evol. 49:729-735(1999).
DR EMBL: AB025323; BAA88476.1; -.
DR InterPro: IPR000886; -.
DR InterPro: IPR001580; -.
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PRODOM: PD001866; -.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 1
```

SQ SEQUENCE 321 AA; 37367 MW; 6E8DFA98D42F7AEF CRC64;

Query Match	68.2%	Score 217;	DB 13;	Length 321;
Best Local Similarity	66.1%	Pred. No. 2.9e-17;		
Matches 39; Conservative	9;	Mismatches 11;	Indels 0;	Gaps 0;

Qy 1 GPGTKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 59
- - - - -
Db 37 GYSTKKVHVILNSKGNHLIKKEVKCKDDETHLYTLIMLYPDQTYEVRVDOEKIESGSL 95

Search completed: October 21, 2001, 03:10:09
Job time: 344 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:05:47 ; Search time 76 Seconds
(without alignments)
223.352 Million cell updates/sec

Title: US-09-807-148-9
Perfect score: 1553
Sequence: 1 GPCFKVHVIFNKGKNVLI.....BEDKREDEEDVPGQAKDEL 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553	100.0	280	21	AA192355
2	1553	100.0	400	21	AA192350
3	1553	100.0	417	10	AA192276
4	1553	100.0	417	20	AA190927
5	1553	100.0	417	21	AA192349
6	1442.5	92.9	401	18	AA111156
7	1022.5	65.8	403	17	AA1904171
8	935.5	60.2	336	12	AA192312
9	856.5	55.2	415	22	AA196341
10	856.5	55.2	415	22	AA196343
11	823	53.0	312	21	AA1924609

12	823	53.0	421	21	AA1924608	Arabidopsis thalia
13	823	53.0	424	21	AA1924607	Arabidopsis thalia
14	816	52.5	312	21	AA1924607	Arabidopsis thalia
15	816	52.5	421	21	AA1924607	Arabidopsis thalia
16	816	52.5	424	21	AA1924607	Arabidopsis thalia
17	801.5	51.6	332	21	AA1924607	Arabidopsis thalia
18	801.5	51.6	441	21	AA1924607	Arabidopsis thalia
19	801.5	51.6	444	21	AA1924607	Arabidopsis thalia
20	702	45.2	122	20	AA1924607	Human ccl1qR binding
21	699	45.0	122	20	AA1924607	Rat ccl1qR binding
22	678	43.7	122	20	AA1924607	Mouse ccl1qR binding
23	624	40.2	593	16	AA1924607	Calnexin sequence.
24	562.5	36.2	385	21	AA1924607	Human secreted pro
25	501.5	32.3	394	21	AA1924607	Arabidopsis thalia
26	501.5	32.3	394	21	AA1924607	Arabidopsis thalia
27	501.5	32.3	530	21	AA1924607	Arabidopsis thalia
28	501.5	32.3	530	21	AA1924607	Arabidopsis thalia
29	501.5	32.3	567	21	AA1924607	Arabidopsis thalia
30	493	31.7	542	22	AA1924607	Castor bean calnex
31	479.5	30.9	394	21	AA1924607	Arabidopsis thalia
32	479.5	30.9	532	21	AA1924607	Arabidopsis thalia
33	479.5	30.9	548	21	AA1924607	Arabidopsis thalia
34	446	28.7	85	17	AA1924607	Flea calreticulin
35	351	22.6	70	21	AA1924607	Human cancer assoc
36	339	21.8	274	21	AA1924607	Arabidopsis thalia
37	318	20.5	60	21	AA1924607	Recombinant human
38	318	20.5	61	21	AA1924607	Recombinant human
39	318	20.5	180	21	AA1924607	Human vasostatin (
40	308.5	19.9	417	21	AA1924607	A. thaliana enviro
41	297	19.1	119	21	AA1924607	Zea mays protein f
42	258	16.6	49	21	AA1924607	Recombinant human
43	246	15.8	235	21	AA1924607	Arabidopsis thalia
44	203	13.1	57	21	AA1924607	Human secreted pro
45	200.5	12.9	515	21	AA1924607	A. thaliana enviro

ALIGNMENTS

RESULT 1

AA192355
ID AA192355 standard; Protein; 280 AA.

XX AA192355;

XX 10-AUG-2000 (first entry)

XX Recombinant delta-120 calreticulin.

XX MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;

XX endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;

XX cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;

XX anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.

OS Synthetic.

XX WC2000020577-A1.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.

XX 06-OCT-1998; 98US-0103438.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Tosato G, Pike SE, Yao L;

XX WPI; 2000-303767/26.

XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,

XX useful for suppressing tumor growth

XX PS Claim 4; Page 86; 99pp; English.

XX PS This sequence comprises recombinant human calreticulin (AA923350) missing the N-terminal 120 amino acids.

XX CC A novel method of inhibiting endothelial cell growth comprises contacting the cells with calreticulin (or its fragments/variants). The fragments of calreticulin causes at least 40% inhibition of angiogenesis, tumor growth and/or endothelial cell growth (claimed). The method may be used for inhibiting angiogenesis in a patient. The angiogenesis is associated with a disease other than a tumor that is associated with neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune inflammation, atherosclerosis, excessive wound repair, retinal neovascularization, macular degeneration, corneal graft rejection, contact lens overwear, Crohn's disease, non-immune inflammation, rheumatoid arthritis, systemic lupus erythematosus, thyroiditis, Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's syndrome, sarcoidosis and primary biliary cirrhosis). The method may also be used for treating/inhibiting tumor growth especially Kaposi's sarcoma (claimed).

XX CC Sequence 280 AA;

Query Match 100.0%; Score 1553; DB 21; Length 280;
Best Local Similarity 100.0%; Pred. No. 7.4e-122;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
DB 1 GPTCKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 60

QY 61 DWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPHIPDPDAKKPEDWDEEM 120
DB 61 ddwdfllppkkikdpdaskpedwderakidpdtkskpedwkphehipdpdakkpedwdeem 120

QY 121 DGEWEPVTONPEYKGEWKPRQIDNPDKGTWTHPEIDNPEYSPDPSIYAYDNFGVLGLD 180
DB 121 dgewepvtonpeykgewkprqidnppdykgtwihpeidnpeyspdpssiayaydnfgvlgld 180

QY 181 LMQVKSgtifdnflitndeyaeefgnwtgvtkaekomkdkodeeorkkeeedkkrk 240
DB 181 lmqvksgtifdnflitndeyaeefgnwtgvtkaekomkdkodeeorkkeeedkkrk 240

QY 241 EEEAEKDEDKDE 280
DB 241 eeeaeekdedkdedeekdedeekdedeekdedeekdedeekdedeekdedeekdedeek 280

RESULT 2
AA923350
ID AA923350 standard; Protein; 400 AA.
AC AA923350;
XX 10-AUG-2000 (first entry)
XX Recombinant human MBP-calreticulin.
KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition; endothelial cell; anti-angiogenic; neuroprotective; antidiabetic; cytostatic; dermalogical; immunosuppressive; anti-inflammatory; hepatic; anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
OS Homo sapiens.
XX WO200020577-A1.
PN 13-APR-2000.
XX 05-OCT-1999; 99WO-US23240.
XX Sjorens syndrome; systemic lupus erythematosus.

PR 06-OCT-1998; 98US-0103438.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Tosato G, Pike SE, Yao L;
XX WPI; 2000-303767/26.
DR N-PSDB; AAA09346, AAA09347.
XX Inhibiting endothelial cell growth and angiogenesis using calreticulin, useful for suppressing tumor growth
XX Claim 4; Page 80-81; 99pp; English.
XX Recombinant human MBP-calreticulin comprises the sequence of human calreticulin (see AA92349) without the 17 N-terminal amino acids.
XX A novel method of inhibiting endothelial cell growth comprises contacting the cells with calreticulin (or its fragments/variants). The fragments of calreticulin causes at least 40% inhibition of angiogenesis, tumor growth and/or endothelial cell growth (claimed). The method may be used for inhibiting angiogenesis in a patient. The angiogenesis is associated with a disease other than a tumor that is associated with neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune inflammation, atherosclerosis, excessive wound repair, retinal neovascularization, macular degeneration, corneal graft rejection, contact lens overwear, Crohn's disease, non-immune inflammation, rheumatoid arthritis, systemic lupus erythematosus, thyroiditis, Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's syndrome, sarcoidosis and primary biliary cirrhosis). The method may also be used for treating/inhibiting tumor growth especially Kaposi's sarcoma (claimed).

XX CC Sequence 400 AA;

Query Match 100.0%; Score 1553; DB 21; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.2e-121;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
DB 121 GPTCKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180

QY 61 DWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPHIPDPDAKKPEDWDEEM 120
DB 181 ddwdfllppkkikdpdaskpedwderakidpdtkskpedwkphehipdpdakkpedwdeem 240

QY 121 DGEWEPVTONPEYKGEWKPRQIDNPDKGTWTHPEIDNPEYSPDPSIYAYDNFGVLGLD 180
DB 241 dgewepvtonpeykgewkprqidnppdykgtwihpeidnpeyspdpssiayaydnfgvlgld 300

QY 181 LMQVKSgtifdnflitndeyaeefgnwtgvtkaekomkdkodeeorkkeeedkkrk 240
DB 301 lmqvksgtifdnflitndeyaeefgnwtgvtkaekomkdkodeeorkkeeedkkrk 360

QY 241 EEEAEKDEDKDE 280
DB 361 eeeaeekdedkdedeekdedeekdedeekdedeekdedeekdedeekdedeekdedeek 400

RESULT 3
AAP92276
ID AAP92276 standard; protein; 417 AA.
XX AAP92276;
XX 23-FEB-1990 (first entry)
XX 60 KD RO (Ro/SSA) antigen.
XX Sjorens syndrome; systemic lupus erythematosus.

OS	Synthetic.	
XX	WO8909273-A.	
PN		
XX		
XX	05-OCT-1989.	
PD		
XX		
XX	22-MAR-1989; 89WO-US01213.	
PF		
XX		
XX	22-MAR-1988; 88US-0171634.	
PR		
XX	(TEXA) UNIV OF TEXAS SYST.	
PA		
XX	Sontheimer RD, Capra JD, McCauliffe DP;	
PI		
XX	WPI; 1989-309537/42.	
XX	N-PSDB; AAP92276.	
DR		
PT	DNA sequences encoding antigenic epitope(s) of Ro 60 kD autoantigen	
PT	- used in immunoassays to detect rheumatic disease	
XX		
PS	Disclosure; Fig 2; 89pp; English.	
XX		
CC	Synthetic peptides corresp. to an epitopic core of Ro antigen are	
CC	expressed recombinantly to detect autoantibodies, for identification	
CC	of autoimmune diseases. These epitopes are AAS 24-36, 23-36, 188-209,	
CC	or 241-255. The peptides may be substd. for ribonucleoprotein particle	
CC	antigens.	
XX		
SQ	Sequence 417 AA;	

XX WO9907406-A1.
XX PN
XX XX
XX 18-FEB-1999.
XX PD
XX XX
XX 12-AUG-1998; 98WO-GB02430.
XX PF
XX XX
XX 12-AUG-1997; 97GB-0016998.
XX PR
XX XX
XX (UYLE-) UNIV LEICESTER.
XX PA
XX Schwaeble W;
XX PI
XX WPI; 1999-180404/15.
XX DR
XX ..
XX Use of a cClqR binding domain - to modulate complement ubiquitin
XX (CUB) functionality.
XX PT
XX PS
XX Disclosure; Page 26-27; 31pp; English.
XX PS
XX This sequence is calreticulin, a homologue of Clq and collectin receptor
XX (cClqR). The invention relates to the use of a cClqR binding domain in a
XX medicament to effect complement ubiquitin (CUB) functionality, and an
XX inhibitor of the cClqR binding domain in a medicament to inhibit CUB
XX functionality. The cClqR binding domain, or its inhibitor, can be used to
XX treat a human or animal body. Particularly an inhibitor is used to treat
XX complement activation involved in the initiation and maintenance of
XX inflammation, for example in myocardial infarction, brain ischaemia
XX (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
XX erythematosus, burns, immune complex nephritis, and to treat amyloid
XX plaques in Alzheimer's disease. The use of cClqR binding domain or
XX inhibitor enables the CUB domain functionality to be modulated using a
XX low molecular weight molecule.
XX CC
XX Sequence 417 AA;
XX SQ


```

RESULT      7
AAW04171
ID   AAW04171 standard; Protein; 403 AA.
XX
XX   AAW04171;
XX
XX   12-DEC-1996   (first entry)
XX
XX   Flea calreticulin PctCal403.
XX
XX   Calreticulin; flea; haematophagous insect; allergic dermatitis;
XX   vaccine; therapy; PctCal403.
XX
XX   Ctenocephalides felis.
XX
XX   WO9628469-AI.
XX
XX   19-SEP-1996.
XX
XX   08-MAR-1996;    96WO-US03133.
XX
XX   09-MAR-1995;    95US-0401509.
XX
XX   (HESK-) HESKA CORP.
XX
XX   Rushlow KE, Stiegler GL;
XX
XX   WPI; 1996-442861/44.
XX
XX   N-PSDB; AAT39516;
XX
XX   N-PSDB; AAT39517.
XX
XX   Haematophagous insect calreticulin protein - used to reduce insect
XX   infestation and desensitise patients to allergic dermatitis
XX
XX   Claim 5; Page 68-69; 86pp; English.
XX
XX   Flea calreticulin protein PctCall589 (AAW04171) is a calcium-binding
XX   protein found in the salivary glands of Ctenocephalides felis.
XX   Its amino acid sequence was deduced from a cDNA clone (AAT39516)
XX   obtd. from a salivary gland cDNA library. Recombinant PctCall589
XX   can be produced in host cells transformed with a vector carrying
XX   calreticulin nucleic acids. Calreticulin alters the blood feeding
XX   behaviour of haematophagous insects and can be administered to an
XX   animal to reduce infestation. It reduces calreticulin activity in
XX   insects, so reducing the insect burden on an animal. Calreticulin
XX   can be used to elicit an immune response, thereby desensitising an
XX   animal to allergic dermatitis caused by fleas, mosquitoes or
XX   Cullcoïdes.
XX
SQ   Sequence      403 AA;

Query Match          65.8%; Score 1022.5; DB 17; Length 403;
Best Local Similarity 68.5%; Pred. No. 1.9e-77;
Matches 185; Conservative 28; Mismatches 48; Indels 9; Gaps

QY       2   PGTKKHVHIFNYGKNVLINCKDKDDFTHLYTLIVRPDNTYEYVKIDNSOVSGSLED 61
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db       139  pgtckkhvifsygknvlInkldrcdddythvytlvvpdntyevlidnekevsgnld 198
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY       62   DWDLPPKKIKDPASKPEDWDERAKIDBPTDSKPEDWDKPEHIPDPDAKKPEDWDEMD 121
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db       199  dweflapkkikdeapkapwderatipndtckpedwdkpehipdpdatkpddwdeemd 258
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY       122  GEWEPPVIQNPEYKGWKPFQIDNPDYKGTWIHPETIDNPEYSPDPSIYAYDNFGVLGLDL 181
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db       259  gewepamidnpaykgewapqldnpykgvwvhpeIdnpeyvpdtqlkyrdeicaigldl 318
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY       182  WQVKSCTIFDNFLITNDENAVAEFGNETWCV--TKAAEQMKRDKQDEEORLKEEEDKKR 239
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db       319  wqvkgacifdnlltdodvyyakiaee---gvkstgegekmmkdagdeerkareae---t 372
         |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY       240  KEEEAEDKED-DEDKDDEDEDEEDEE 268

```


PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.0%; Score 823; DB 21; Length 312;
Best Local Similarity 54.0%; Pred. No. 6e-61;
Matches 155; Conservative 36; Mismatches 86; Indels 10; Gaps 5;

QY 1 GPGTKKHVIFNYKGNVLINIKDKDEFFHLYTLIVRPONTYEVKIDNSQVESGSLE 60
Db 29 gystkxvhailltynanhlkkdvpcetdqlthvyflirpdatsilldnvekqtgsly 88
QY 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWK-PEHIPDPDAKKPEDWDEE 119
Db 89 sdwllppkkikdpsakpedwdeqeyisdpedkpgdyddipkeipdtdskkpedwde 148
QY 120 MDGEWEPPIQNPYKGEWKPROIDNPDKYWIHPEIDNPESYSPDIYAYDNFVGL 179
Db 149 edgewtaptipnpeymgewkpgkikpnkkgweaplidnpgdfkdpelyvfpkkyvg 208
QY 180 DLWQVKSFTIFDNFLITNDEAYAEFEFGNETGWTKAAEKQMKODEEORLKEEEDKKR 239
Db 209 elwqvksgsifdnvlicddpdyakkladetwglkdaeka---atdeaaekneeeskda 265
QY 240 KEEEAEDK-EDDE---DKDEDEDEDEDEDEDEE--DVPQAKDEL 280
Db 266 paesdaedepeddegddsdsseskaetksedsetsektdatahd 312

RESULT 12

AAG24608

ID AAG24608 standard; Protein; 421 AA.

XX AC AAG24608;
XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28350.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

OS Arabidopsis thaliana.

XX PN EP1033405-A2.
XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	06-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	09-AUG-1999;	99US-0147935.
PR	14-JUN-1999;	99US-0139119.	PR	10-AUG-1999;	99US-0148171.
PR	16-JUN-1999;	99US-0139452.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139453.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139454.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139455.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139456.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139457.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139462.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139463.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139750.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139899.	PR	27-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	99US-0140353.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140354.	PR	31-AUG-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140695.	PR	01-SEP-1999;	99US-0151930.
PR	28-JUN-1999;	99US-0140823.	PR	07-SEP-1999;	99US-0152363.
PR	29-JUN-1999;	99US-0140981.	PR	07-SEP-1999;	99US-0153070.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143342.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140693.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.


```
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 52.5%; Score 816; DB 21; Length 421;
Best Local Similarity 53.7%; Pred. No. 3.4e-60;
Matches 154; Conservative 36; Mismatches 87; Indels 10; Gaps 5;

QY 1 GPC TKKVHVIFNYKGNVLINCKDDETHLYTLIVRPDNTYEVKIDNSOVESSSLE 60
DB 138 gystkkvhaillyneanhlkkdvpctdqlthvylflirpdatsilldnvektgsgly 197
QY 61 DDWDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDK-PEHIPDPDAKKPEDWDEE 119
DB 198 sdwllppkklkdpakpewdeqeyisdpedkpgdyddlpkeipdtcdskkpedwdde 257
QY 120 MDGEWEPPIQNPPEYKGEWKPRQIDNPDPYKGTWHPIDNPENPEYSPDPSPYAYDNFVGL 179
DB 258 edgewtaptipnpeymgewkpgkqknpn/ykgweapllidnpdfkdpelyvfpkkyvgl 317
QY 180 DLWQVKSgtIFDNFLITNDEAYAEFGNETWGTWKAQKQKODEORLKEEEDKKR 239
DB 318 elwqvksgsifdnvllccddpdyakkladewtqklkdaeka---afdeakknneeskda 374
QY 240 KEEEAEDK-EDDE---DKDEDEDEDEKEDEEE---DVPQAKDEL 280
DB 375 paesdaedepdeddegddsdskaeetksvdseetsekdahdel 421

Search completed: October 21, 2001, 03:05:48
Job time: 343 sec
```



```

Db 199 TDWLLPPKPKIDPEAKKPDWDEKXYIPDPEDKKPEGYDDIPKEIPDPDAKPEDWDDE 258
QY 120 MCGEWEPPVIONPEYKGEWKPQROINDPNKYGTWHPEDINPEYSPDPISYAYONFGVLGL 179
Db 259 EDGEWTAPTIANPEYKGPWKPKKIKNPYKGGKAPMINDPDFKDDPEYVYNLYVGI 318
QY 180 DLQVKSGLTFDNLITNDAYAEERFNETGWVTKAAEQKMDKODEQRLKEEEDKKR 239
Db 319 ELQVKSGLTFDNLITNDPEYAKLAETGWGNKDAEKA---AFEEAEKKKEEESKDD 375
QY 240 KEEEAEDKEDDEDDEDEDEK---EDEDVPGQAKDEL 280
Db 376 PADSDADEDDDDADDTEGEDDGESKDAEADSDV---HDEL 415

```

```

RESULT 2
US-08-296-362-2
; Sequence 2, Application US/08296362
; Patent No. 5691306
; GENERAL INFORMATION:
; APPLICANT: Bergeron, John J.M.
; APPLICANT: Thomas, David Y.
; APPLICANT: Wada, Ikao
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; TITLE OF INVENTION: PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,362
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deehr, Manya S.
; REGISTRATION NUMBER: 37,120
; REFERENCE/DOCKET NUMBER: 690066.401C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-296-362-2

```

```

Query Match 40.2%; Score 624; DB 1; Length 593;
Best Local Similarity 33.7%; Pred. No. 1.1e-48;
Matches 129; Conservative 58; Mismatches 86; Indels 110; Gaps 8;
QY 6 KVHIVFYKGNKVLINKDIRCK-----DDETHLYTLIVRPDNTYEVKIDNSQVE 55
Db 200 KLFHFRHNKPTGVTEYKHAKRPADLKYTFTRKTHLYTLILNPDNSFEILVDSIYN 259
QY 56 SGSLEDDMDFLPP-----KKIKPDASKPDWDERAKIDDPDTSKPDW----- 99
Db 260 SGNLLN-DMTTPVNSRIEDPEQKPEDWDERKIPDPDAVKPDWDDWEDAPAKIPDEE 317

```

```

QY 100 -----DKPHIIPDPDAKPEDWDEEMDGWEP-----PVION 131
Db 318 ATKPDGWLDDPEYVPPDAEKPEDWDEMDGWEAPQIANPCESAPGCGVQWRPMIDN 377
QY 132 PEYKGEWKPQROINDPNKYGTWHPEDINPEYSPDPISYAYDNFGVLGLDLMQVKSGLTFD 191
Db 378 PNYKWKGPMPIDNPYQGIWKPKIPNPDFFEDLEPFKMTPTFSAIGLELWSMTSDIFFD 437
QY 192 NFLITNDAYAEERFNETGWVTKAAE-----KOMKDKQDS----- 226
Db 438 NFIVCGRVRRVDDWANDGWLKKAADGAAPGVVGOMIEAAEERPMWLWVYVITVALPVF 497
QY 227 -----EORLKEEEDK-----KRKEEBAEDKEDDEKDED 257
Db 498 LVISFCCSGKQSSPVEYKKTDAQPDVKEEKEEKEEKEEKEEKEEKEEKEEKEEKEE 557
QY 258 EDEEDKEDEEDVPGQAKDEL 280
Db 558 EEDGTSQAEEDDRKPKAEDEI 580
RESULT 3
US-08-675-816-6
; Sequence 6, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND US
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave., Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)-622-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-816-6

```

```

Query Match 31.7%; Score 493; DB 4; Length 542;
Best Local Similarity 29.8%; Pred. No. 7.3e-37;
Matches 117; Conservative 41; Mismatches 113; Indels 122; Gaps 10;
QY 1 GP-----GTTKVVHVFNYKG--KNVLINKDIR-----CKDDEFTHTLYTLIVRPDNTYEVKID 50
Db 143 GPKDKCATKVKVFLKHKNPCKGEYIEHHLKYPSPVSKLTHVYTAILKPDNELILVD 202
QY 51 NSQVESGSLEDDMDFLPP-----KKIKPDASKPDWDERAKIDDPDTSKPDW----- 99
Db 203 GEEKKANFLSSDDFEPPLVPARTIPDPDDKKPEDWDERAKIPDPNAPVDPDWDDEAPME 262

```

Query Match 11.9%; Score 184.5; DB 2; Length 714;

	Qy	38	I VRPDNTYEVKINSQVSGSLDDWD-----FLPPKKIKDPDASKPDPWDERAKID	89
			: : : :	
	Db	526	VPRPAFVHYDKKEASDVET--SLESDDSDSVIIVEGLPLPLPPPESGATP---PPIAPTg	581
	Qy	90	DPTDSKP-EDWDKPEHTPDPAKDAPDWOBEMDGMEPPVIQNPEYKGWKPRQINDPDY	148
			: :	
	Db	582	PPTASPVPVPAKEEPEELPAAGPLPPP-----ppp-----	611
	Qy	149	KGTWIHPINPNEXSPDPSIIAYDNFG----VLGLDLMOVKSGTIFDFNFLTINDEAYAE	203
			: :	
	Db	612	-----PPVPGPVXLPPQLPVLPSTPGGGPALERDL-----tVIN--INSDBEEEEE	658
	Qy	204	EFGNETGVTTAAEKOMKQDOEQRLKKEEDEKKRKEEEADEKDEDKDEDEE--dD	260
			: :: : ::: :: :: :: :: :: :: :: :: ::	
	Db	659	EGEEEEEEEEEEEEEEEEEEEEEFEEEEDEEYfEEEEEEEPfEEEEFEeGEUle	718
	Qy	261	FEDKEEDEED 271	
			: :	
	Db	719	EEEEEBEERE 729	

```

RESULT      8
US-08-574-959A-7
: Sequence 7, Application US/08574959A
: Patent No. 5962224
: GENERAL INFORMATION:
: APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
: APPLICANT: and Jack L. Strominger
: TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
: TITLE OF INVENTION: AND USES THEREFOR
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSES: LAHIVE & COCKFIELD
: STREET: 60 State Street, Suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/574,959A
: FILING DATE: 19-DEC-95
: ATTORNEY/AGENT INFORMATION:
: NAME: Mandragouras, Amy E.
: REGISTRATION NUMBER: 36,207
: REFERENCE/DOCKET NUMBER: DFN-008
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1135 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-574-959A-7

```

Query Match	10.5%	Score 163;	DB 2;	Length 1135;
Best Local Similarity	26.3%	Pred. No. 1.6e-06;		
Matches 66;	Conservative	36;	Mismatches 85;	Indels 64;
				Gaps 10;

[illegible]

Qy	149	KGTWIIHEIDNPEYSPDPSTAYDNFG-----VLGLDLQVKGSGTTFDNLFLINDAYAE	20303
Db	842	-----PPVPGCVLPPQLVPDPTCGGGPPALEED-----TWIN--INSDEEEEE	8988
Qy	204	EFGNETGVTKAAEKQMKDKODEBQRUKBEEEDKKKEEEAEKDEDEKDEDE---D	2600
Db	889	EGEEEEEEEEEEEEEEEEEEEDFESEEDFESEEEFYFESEEEEEFEFEFESEGE	9488
Qy	261	EEDKEDEDEED	271
Db	949	EEEEDEEEEE	959

RESULT 9
 US-09-022-983-5
 : Sequence 5, Application US/09022983
 : Patent No. 6159731
 : GENERAL INFORMATION:
 : APPLICANT: Yang, Xiaolu
 : APPLICANT: Khosravi-Far, Roya
 : APPLICANT: Chang, Howard Y.
 : APPLICANT: Baltimore, David
 : TITLE OF INVENTION: DAXX, A NOVEL FAS-BINDING
 : TITLE OF INVENTION: PROTEIN THAT ACTIVATES JNK AND APOPTOSIS
 : NUMBER OF SEQUENCES: 5
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 : STREET: 600 Atlantic Avenue
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02210
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: Fastseq for Windows
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/022,983
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/037,919
 : FILING DATE: 12-FEB-1997
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/051,753
 : FILING DATE: 26-JUN-1997
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Van Amsterdam, John R.
 : REGISTRATION NUMBER: 40,212
 : REFERENCE/DOCKET NUMBER: M06566/7036
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 617-720-3500
 : TELEFAX: 617-720-2441
 : INFORMATION FOR SEQ ID NO: 5:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 740 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : FRAGMENT TYPE: internal
 : US-09-022-983-5

```

Query Match      9.7%  Score 151;  DB 4;  Length 740;
Best Local Similarity 28.4%;  Pred. No. 1.1e-05;
Matches 44;  Conservative 23;  Mismatches 40;  Indels 48;  Gaps 5;
QY 128 VIQPEYKGEWKPRQIDNPDKGTWIHPHPEINPEVSPD----PSIYAYDNFCVLGLDLWQ 183
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 381 MLOPKSEGEKKRR---ARLQGTSSH-SADTPKASLDGEGPS-----SADTPKASLDGEGPS 420

```

Qy	184	VKSCTIFDNFLITNDEAYAEFNETGWVTKAAEKMKDKQDEEQRLKEEEDKKRKEE	243
Db	421	-----GMA5QGCP5ASRAETDDDEDES--DEEEEEEEEE	456
Qy	244	EAEKDEDDKDEDEDEKDEEDEDVPGAKD	278
Db	457	EAYDSEEDLEOMGOEDEDEDEEEEAAGKD	491

```

RESULT 10
US-08-935-855-20
; Sequence 20, Application US/08935855
; Patent No. 6066485
; GENERAL INFORMATION:
; APPLICANT: Guthridge, Mark
; APPLICANT: Basílico, Claudio
; TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
; TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,855
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-935-855-20

```

[illegible]

```

QY 192 NFLTINDEATAE-----EFGNETWGTGK-----AAEKQMKDKODEORLKEE 234
Db 216 -----SNSRTEAQCVGPGIPTGEAGPSCSSASDKLPRVAKSKFFEDSEDEAE 271
QY 235 EDKRRKEE-----AEDKDEDEKDEDEDEDEDEE-DVPG-QAKDE 279
Db 272 EDEECSEEDGYSSEAEANEDEDETEAEAEDEDEDEDEDEEMVPGWEGKE 322

RESULT 11
US-08-741-134-6
: Sequence 6, Application US/08741134
: Patent No. 5861498
: GENERAL INFORMATION:
: APPLICANT: Litwack, Gerald
: APPLICANT: Alnemri, Emad S.
: APPLICANT: Fernandes-Alnemri, Teresa
: TITLE OF INVENTION: IMMUNOPHILIN FKBp46 AND COMPOSITIONS FOR MAKING
: TITLE OF INVENTION: AND
: TITLE OF INVENTION: METHODS OF USING THE SAME
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5861498ris
: STREET: One Liberty Place - 46th floor
: CITY: Philadelphia
: STATE: Pennsylvania
: COUNTRY: USA
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows 3.11
: SOFTWARE: WordPerfect for Windows 6.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/741,134
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/007,163
: FILING DATE: 01-NOV-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: DeLuca, Mark
: REGISTRATION NUMBER: 33,229
: REFERENCE/DOCKET NUMBER: TJU-2090
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-568-3100
: TELEFAX: 215-568-3439
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 411 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PS-08-741-134-6

```

```

Query Match          9.6%; Score 149; DB 2; Length 411;
Best Local Similarity 21.2%; Pred. No. 7.8e-06;
Matches 49; Conservative 53; Mismatches 82; Indels 48; Gaps 5;

QY 50 DNSOVESGSLDDWDFLPPKKIKDPDASKPDWDERAKIDDPQSKPDWDPKPEHIP-DP 108
    | : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 72 DEEDIDSESEEEEEETQKKK---SKGKAESSEDEDEDEDEDEFOEYVLLTSLP 127
    | : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 109 DAKKPDWDEMDCEWPPPIQNPEYKGWKPRQIDPNPYKWIHPIDNPESYDPDSI 168
    | : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 128 EAQYQOSLDLTIPEEEVQIVTGSYA-----ISLGNVYKHPDTP----- 169

QY 169 YAYDNFVGLDLDMQVKSGETFDNFLTNDPEAYAEFGNETGWGTKAAEKOMKQDEEQ 228
    | : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 170 -----MGVGED-----EDEDADIYDSYDITPDEETIGQDMDD-- 205
    | : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |

```



```

111 DGKVMISPKRAEEVAFFAKMDLHVEYTTKEIFKRNFKD-WRKEMTNEEKNIITNSKCD 169
      . : | : | : | | | | | : | : | : | : | : | : |
217 EKQWKD---KODFEORIKKFEEDKKKKKFEAFDKE 249
      . : | : | : | | | | | : | : | : | : | : | : |

```

Search completed: October 21, 2001, 03:06:42
Job time: 332 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:07:45 ; Search time 52.92 Seconds
(without alignments)
403.040 Million cell updates/sec

Title: US-09-807-148-9
Perfect score: 1553
Sequence: 1 GPGTKKHVIFNYKGNVLI.....EEDKEEDBEDVPGQAKDEL 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553	100.0	417	1 A37047	calreticulin precu
2	1487.5	95.8	416	1 S06763	calreticulin precu
3	1476.5	95.1	416	2 JH0819	calreticulin precu
4	1457.5	93.9	418	1 A34154	calreticulin precu
5	1442	92.9	400	2 S43376	calreticulin, brai
6	1432	92.2	421	2 S36799	calreticulin precu
7	1280.5	82.5	411	2 S29129	calreticulin precu
8	1249.5	80.5	384	2 S29130	calreticulin (clon
9	1241.5	79.9	419	2 S71143	calreticulin precu
10	1164	75.0	405	1 JH0795	calreticulin precu
11	1031.5	66.4	406	2 A56637	calreticulin homol
12	991.5	63.8	395	2 S25851	calreticulin precu
13	944.5	60.8	336	2 A32507	calreticulin precu
14	856.5	55.2	415	2 T10172	41k larval antigen
15	844	54.3	421	2 S58170	calreticulin - cas
16	818.5	52.7	416	2 T16968	calreticulin precu
17	817.5	52.6	399	2 T03691	calreticulin call
18	817.5	52.6	393	1 A48573	calreticulin - com
19	814.5	52.4	416	2 T14554	calreticulin autoa
20	813	52.4	422	2 T07841	calreticulin - bee
21	812	52.3	412	2 T05703	probable calreticu
22	812	52.3	415	2 T05705	calreticulin - bar
23	810.5	52.2	435	2 C96605	calreticulin (Crtl
24	801.5	51.6	444	2 H86224	hypothetical prote
25	643	41.4	622	2 S71342	calnexin precursor
26	633	40.8	591	2 B54354	calnexin precursor
27	629	40.5	591	2 C54354	calnexin precursor
28	628	40.4	592	2 I53260	calnexin - human
29	628	40.4	592	2 A46673	calnexin precursor

30	624	40.2	593	1 A37273	calnexin precursor
31	587	37.8	356	2 A46164	calnexin - human (
32	555	35.7	611	2 A54086	calnexin-t - mouse
33	554	35.7	611	2 A53418	calmeglin precursor
34	539.5	34.7	582	2 A46637	calnexin homolog S
35	531	34.2	546	2 T06415	calnexin - soybean
36	501.5	32.3	428	2 T03251	calnexin - maize (
37	501.5	32.3	530	2 JN0597	calnexin-like prot
38	477.5	30.7	532	2 T49873	calnexin homolog -
39	472.5	30.4	560	2 S56142	calcium-binding pr
40	472.5	30.4	619	2 S40938	hypothetical prote
41	458.5	29.5	297	2 S70552	calnexin homolog C
42	451.5	29.1	540	2 T10852	probable calnexin
43	299.5	19.3	502	2 S29347	calnexin homolog Y
44	208.5	13.4	792	2 T42963	hypothetical prote
45	193	12.4	798	2 T33022	hypothetical prote

ALIGNMENTS

RESULT 1

A37047
calreticulin precursor - human
N:Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding pr
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Feb-2000
C:Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075
R:McCaulliffe, D.P.; Yang, Y.S.; Willson, J.; Sontheimer, R.D.; Capra, J.D.
J. Biol. Chem. 267, 2557-2562, 1992
A:Title: The 5'-flanking region of the human calreticulin gene shares homology with t
A:Reference number: A42330; MUID:92129342
A:Accession: A42330
A:Molecule type: DNA
A:Residues: 1-417 <MC2>
A>Note: sequence extracted from NCBI backbone (NCBIN:78537, NCBIP:78536)
R:McCaulliffe, D.P.; Lux, F.A.; Lieu, T.S.; Sanz, I.; Hanke, J.; Newkirk, M.M.; Bachin
J. Clin. Invest. 85, 1379-1391, 1990
A:Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/
A:Reference number: A37047; MUID:90237213
A:Accession: A37047
A:Molecule type: mRNA
A:Residues: 1-417 <MCC>
A:Cross-references: GB:M32294; NID:9337486; PIDN:AAA36582.1; PID:g337487
A>Note: the authors translated the codon GTA for residue 349 as Tyr
R:Rokeach, L.A.; Haselby, J.A.; Meilof, J.F.; Smeenk, R.J.; Unnasch, T.R.; Greene, B.
J. Immunol. 147, 3031-3039, 1991
A:Title: Characterization of the autoantigen calreticulin.
A:Reference number: A46452; MUID:92013129
A:Accession: A46452
A:Molecule type: mRNA
A:Residues: 1-417 <ROK>
A:Cross-references: GB:M84739; NID:g179881; PIDN:AAA51916.1; PID:g179882
A>Note: sequence extracted from NCBI backbone (NCBIN:60749, NCBIP:60750)
R:Lieu, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.
J. Clin. Invest. 82, 96-101, 1988
A:Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence
A:Reference number: A28812; MUID:88273610
A:Accession: A28812
A:Molecule type: protein
A:Residues: 18-41 <LIE>
A>Note: 18-Ala was also found
R:Dupuis, M.; Schaefer, E.; Krause, K.H.; Tschopp, J.
J. Exp. Med. 177, 1-7, 1993
A:Title: The calcium-binding protein calreticulin is a major constituent of lytic gra
A:Reference number: PH1525; MUID:93115648
A:Accession: PH1525
A:Molecule type: protein
A:Residues: 18-27 <DUP>
A:Experimental source: LAK cell
R:Roijani, M.V.; Finlay, B.B.; Gray, V.; Dedhar, S.
Biochemistry 30, 9859-9866, 1991
A:Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (c

RESULT 5

C:Keywords: calcium binding; glycoprotein
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-421/Product: calreticulin, brain isoform 2 #status predicted <MAT>
F:418-421/Region: endoplasmic reticulum retention signal
F:141-167/Disulfide bonds: #status predicted
F:183/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.2%; Score 1432; DB 2; Length 421;
Best Local Similarity 92.2%; Pred. No. 5.6e-80;
Matches 259; Conservative 12; Mismatches 8; Indels 2; Gaps 2;

QY 1 GPGETKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEYKIDNSQVESGSLE 60
|||
Db 142 GPGETKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEYKIDNSQVESGSLE 201
|||
QY 61 DWDWFLPPKKIKDPDAASKPEDWDRAKIDDPTDSKPEDWKPEHIPODDAKKPEDWDEEM 120
|||
Db 202 DWDWFLPPKKIKDPDAASKPEDWDRAKIDDPTDSKPEDWKPEHIPODDAKKPEDWDEEM 261
|||
QY 121 DGWEPPVIONPEYKGGMKPRQINDPDYKGTWIHPEDINPEYSPDPSTIAYADNFGVLGLD 180
|||
Db 262 DGWEPPVIONPEYKGGMKPRQINDPEYKGIWIHPEDINPEYSPDSNIYAYENAVLGLD 321
|||
QY 181 LWQVKSGTIFDNFLTINDEAYAEFEGNETGVTGVTAAEKQMKDKQDEEORLEEDEDKKRK 240
|||
Db 322 LWQVKSGTIFDNFLTINDEAYAEFEGNETGVTGVTAAEKQMKDKQDEEORLEEHEEEKGK 381
|||
QY 241 EEEBAEDKEDDEDKDEBDEEDEKEDEEED-VPGQAQDEL 280
|||||
Db 382 EEEBAE-KDDDEDKDEDEDEDEKEEEEEEDAAAAAQADEL 421
|||||

RESULT 7
S29129
C:Title: Identification of calreticulin isoforms in the central nervous system.
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: S29129
R:Reves, S.; Zorzato, F.; Pozzan, T.
Biochem. J. 287, 579-581, 1992
A:Reference number: S29129; MUID:93074997
A:Accession: S29129
A:Molecule type: mRNA
A:Residues: 1-411 <REA>
A:Cross-references: EMBU:X67597; NID:g64608; PIDN:CAA47866.l; PID:g64609
C:Superfamily: calreticulins
C:Keywords: glycoprotein
F:1-12/Domain: signal sequence (fragment) #status predicted <SIG>
F:13-411/Product: calreticulin #status predicted <MAT>
F:408-414/Region: endoplasmic reticulum retention signal
F:339/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.5%; Score 1280.5; DB 2; Length 411;
Best Local Similarity 80.8%; Pred. No. 8.3e-71;
Matches 225; Conservative 31; Mismatches 22; Indels 3; Gaps 2;

QY 1 GPGETKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEYKIDNSQVESGSLE 60
|||
Db 133 GPPTKKVHVIFYKKNQLINKDIRCKDDSFTHLYTLIVRPDNTYEYKIDNSKVESGSLE 192
|||
QY 61 DWDWFLPPKKIKDPDAASKPEDWDRAKIDDPTDSKPEDWKPEHIPODDAKKPEDWDEEM 120
|||
Db 193 DWDWFLPPKKIKDPDAASKPEDWRPKIDDPEDKKPEDWKPEHIPODDAVKPEDWDEEM 252
|||
QY 121 DGWEPPVIONPEYKGGMKPRQINDPDYKGTWIHPEDINPEYSPDPSTIAYADNFGVLGLD 180
|||
Db 253 DGWEPPVIONPEYKGGMKPRQINDPDYKGTWIHPEDINPEYTPDITLYSYDFGVGLD 312
|||
QY 181 LWQVKSGTIFDNFLTINDEAYAEFEGNETGVTGVTAAEKQMKDKQDEEORLEEDEDKKRK 240
|||

C;Superfamily: calreticulin
C;Keywords: calcium binding; endoplasmic reticulum
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-419/Product: calreticulin #status predicted <WAT>
F;205-213/Region: nuclear location signal
F;415-418/Region: endoplasmic reticulum retention signal

```
Query Match      79.9%; Score 1241.5; DB 2;  
Best Local Similarity 76.9%; Pred. No. 1.9e-68;  
Matches 216; Conservative 34; Mismatches 30; I  
  
Qy 1 GPGTKKHVHIENYKGKNVLINRKCDDEETHLYTLIVRPDNTVE  
||| |  
||| | : : : : : : : : : : : : : : : : : : : : : : :  
  
Db 139 GPPTKKHVHIENYKGKNLIQINKDIRSKADVSHLYTLIVRPDNTVE
```

[illegible]

RESULT 10
JH0795
calreticulin precursor - California sea hare

C/Alternate names: protein 407
C/Species: Aplysia californica (California sea hare)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_update_date 10-Sep-1999
C/Accession: JH0795; B31409; F60977
R/Kennedy, T.E.; Kuhl, D.; Barzilai, A.; Sweatt, J.D.; Kneafsey, P.A.
Neuron 9, 1013-1024, 1992
A>Title: Long-term sensitization training in aplysia leads to persistent changes in cyclic AMP levels
A/Reference number: JH0795; PMID:9308937
A/Accession: JH0795
A/Molecule type: mRNA
A/Residues: 1-405 <KEND>
A/Cross-references: GB:S51239; NID:g252053; PIDN:AAB24566
R/Experimental source: abdominal ganglion and anal nerve
R/Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.
Proc. Natl. Acad. Sci. U.S.A. 85, 7008-7012, 1988
A>Title: Sequencing of proteins from two-dimensional gel electrophoresis

A:Reference number: A94207; MUID:88320566
A:Accession: B31409
A:Molecule ttype: protein

A: Molecule type: protein
 A: Residues: 'X', 17-28, 'X', 30-31 <SWE>
 C: Superfamily: calreticulin
 C: Keywords: calcium binding; endoplasmic reticulum
 F: 1-15/Domain: signal sequence #status predicted <SIG>
 F: 16-405/Product: calreticulin #status experimental <NAT>
 F: 402-405/Region: endoplasmic reticulum retention signal

Matches	204: Conservative	29: Mismatches	39: Indels	8: Gaps	2:
Qy	1	GPGTKVHVIFNYGKKNVLINKD	IRCKDDSFTHLYTLIVRPDNTYEVKIDNSQVESGSL	60	
Db	134	GPGTKVHVIFNYGKKNLLVKD	IRCKDDVFSLHYTLIVRPDNTYEVKIDNEAESGDLE	193	
Qy	61	DDWDFLPKKIKDPDASKPEDWDERAKID	DDTDSKPEDWDKPEHIPDPAKKEPDWDEM	120	
Db	194	ADWDFLFAKTI	PPDPAKPPDWDREKIDDPDTKPEDWDKPEHIPDPEAKKPPDWDDEM	253	
Qy	121	DGEWEPVIONPEYKGEMWKPRQ	IDNPDKYGTWTHPEIDNPPEYSPDPSIAYIDNFVGLGLD	180	
Db	254	DGEWEPPIIDNP	EYKGEMWKQVDNPYKGVKWHPEIDNPPEYADDKLYSFADFGAIGFD	313	
Qy	181	LWQVKGSTIFIDNLFITNDWEAYAE	EEFGNETGWGTVKAAEKOMKDKODEORLKEEEDKKR	240	
Db	314	LWQVKAGTIFDNLVITSDVEYAE	EEFGNETGWGTKDKPEKKMKDAQDEEDRKAREEEKKRK	373	
Qy	241	EEEAEDKEDDEDKDEDEDEE	DEEDVPQGAQDEL	280	
Db	374	EEEDA-----NKD-DEEEAE	EEEEEDDDAAPEKDEL	405	

RESULT 11
A56637
calreticulin homolog precursor - fruit fly (*Drosophila melanogaster*)
N;Alternate names: Ro/SS-A autoantigen/calreticulin homolog
C;Species: *Drosophila melanogaster*
C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 13-Aug-1999
C;Accession: A56637; A37158
R;Smith, M.J.

DNA Seq. 3, 247-250, 1992
 A:Title: Nucleotide sequence of a *Drosophila melanogaster* gene encoding a calreticulin h
 A:Reference number: A56637; MUID:93208374
 A:Accession: A56637
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-406 <SMI>
 A:Cross-references: GB:X64461; NID:g7685; PIDN:CAA45791.1; PID:g7686
 A:Note: sequence extracted from NCBI backbone (NCBIN:128274, NCBIPI:128275)
 R:McCaulliffe, D.P.; Zappi, E.; Lieu, T.S.; Michalek, M.; Sontheimer, R.D.; Capra, J.D.
 J. Clin. Invest. 86, 332-335, 1990
 A:Title: A human Ro/SS-A autoantigen is the homologue of calreticulin and is highly hom
 A:Reference number: A37158; MUID:90307981
 A:Accession: A37158
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: DNA
 A:Residues: 91-105, 'A', 107, 109-124; 182-183, 'L', 185-220 <MCC>
 C:Genetics:
 A:Gene: FlyBase:Crc
 A:Cross-references: FlyBase:FBgn0005585
 A:Introns: 65/1; 222/3
 C:Superfamily: calreticulin
 C:Keywords: calcium binding; endoplasmic reticulum
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:403-406/Region: endoplasmic reticulum retention signal

```
Query Match          66.4%; Score 1031.5; DB 2; Length 405;
Best Local Similarity 67.3%; Pred. No. 9.9e-56;
Matches 181; Conservative 36; Mismatches 51; Indels 1; Gaps 1;
```

QY	1	GP	G	T	K	K	V	H	V	I	N	T	Y	N	T	K	K	V	N	L	I	N	K	D	I	R	C	K	D	D	E	F	T	H	L	T	I	V	R	P	D	N	T	Y	E	V	K	I	D	N	S	Q	V	E	S	G	S	L	E		60
Db	138	GP	G	T	K	K	V	H	V	I	N	T	Y	N	T	K	K	V	N	L	I	N	K	D	I	R	C	K	D	D	V	T	H	F	T	L	I	V	R	P	D	N	T	Y	E	V	L	I	D	N	E	K	V	E	S	G	N	L	E		197
QY	61	D	N	D	W	F	L	P	P	K	I	K	D	P	A	S	K	P	E	D	W	E	R	A	K	I	D	D	T	S	K	P	E	D	W	K	P	E	H	I	P	D	P	O	A	K	P	E	D	W	D	E	M		120						
Db	198	D	N	D	F	L	A	P	K	I	K	D	P	A	T	K	P	E	D	W	D	R	A	T	I	P	D	D	K	P	E	D	W	K	P	E	H	I	P	D	P	O	A	K	P	E	D	W	D	E	M		257								
QY	121	D	G	E	W	E	P	P	I	O	N	P	E	Y	K	E	W	P	R	O	I	D	N	P	D	Y	K	C	T	W	I	H	P	E	I	D	N	P	E	S	P	D	S	I	Y	A	V	D	N	F	G	V	I	L	D		180				
Db	258	D	G	E	W	E	P	P	I	D	N	P	E	Y	K	E	W	P	R	O	I	D	N	P	D	Y	K	C	T	W	I	H	P	E	I	D	N	P	E	S	P	D	S	I	Y	A	V	D	N	F	G	V	I	L	D		317				

41K larval antigen - nematode (*Onchocerca volvulus*) (fragment)
C; Species: *Onchocerca volvulus*

Qy	181	LWQVKS	TIFNFI	LITND	EAYAEF	GNETG	VTKAAE	KOMK	DOKD	OEEQR	LKEEBED	KRRK	240
					:	:	:	:	:	:	:	:	
					:	:	:	:	:	:	:	:	
Db	318	LWQVKS	TIFNVI	LITDD	VELAAR	AKAAAF	VKN-TQ	AEGKKM	EAQO	DEVQR	KDEEEKAS	376	
					:	:	:	:	:	:	:	:	
					:	:	:	:	:	:	:	:	
Qy	241	EEEEAE	KDEDK	DEDEE	DEEDEE	DEE	DEE	DEE	DEE	DEE	DEE	269	
					:	:	:	:	:	:	:		
					:	:	:	:	:	:	:		
Db	377	DKDDE	DEDD	DEEK	DESKOD	KDOSEH	DE	405					

RESULT 12

S25851
calreticulin precursor - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 06-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S25851; T33996
R:Smith, M.J.
DNA Seq. 2, 235-240, 1992
A:Title: A *C. elegans* gene encodes a protein homologous to mammalian calreticulin.
A:Reference number: S25851; MUID: 92329978

A:Accession: S25851
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <SMI>
A:Cross-references: EMBL:X59589; NID:g6693; PIDN:CAA42159.1; PID:g6694
R:Bauer, C.; Courtney, L.; Laplant, Y.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of *C. elegans* cosmid Y38A10A.
A:Reference number: Z21453
A:Accession: T33996
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-395 <BAU>
A:Cross-references: EMBL:AF125963; PIDN:AAJ14746.1; GSPDB:GN00023; CESP:Y38A10A.5
A:Experimental source: strain Bristol N2; clone Y38A10A
C:Genetics:
A:Gene: CESP:Y38A10A.5
A:Map position: 5
A:Introns: 107/3; 315/3
C:Superfamily: calreticulin
F:1-15/Domain: signal sequence #status predicted <SIG>
F:392-395/Region: endoplasmic reticulum retention signal

Query Match	63.8%	Score	991.5	DB 2	Length	395
Best Local Similarity	66.0%	Pred. No.	2.6e-53			
Matches	173	Conservative	35	Mismatches	53	Indels
						1
						Gaps
						1

[illegible]

RESULT	13
A22507	

C; Superfamily: calreticulin

RESULT 15
S58170

C;Accession: S58170; S49818
R;Dresselhaus, T.; Hagel, C.; Loerz, H.; Kranz, E.
submitted to the EMBL Data Library, July 1995
A;Description: Isolation of a cDNA encoding Calreticulin from in vitro zygotes
A;Reference number: S58170
A;Accession: S58170
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-421 <DRE>
A;Cross-references: EMBL:X89813; NID:g927571; PIDN:CAA61939.1; PID:g927572
R;Napier, R.M.; Trueman, S.; Henderson, J.; Boyce, J.M.; Hawes, C.R.; Frick
submitted to the EMBL Data Library, November 1994
A;Description: Purification and sequencing of calreticulin from maize and
A;Reference number: S49818

Query Match	Score 844;	DB 2;	Length 421;
Best Local Similarity	54.3%		
Best Local Similarity	53.7%	Pred. NO.	2.4e-44
Matches 154.	Conservative	46.	Matches 71.
			Table 2
			16.
			Case

Qy	180	DLWGVKSGTIFDNLINDEAYAEFQNETWGTAKAEKQMDQKDEBQRLKEEEEDKKR	239
Db	324	ELWGVKSGTIFDNLIIIDPALAKTFAETWGRHKEAKAFADEAK-----KKEEEDA	379
Qy	240	-----KKEEAEADKEDEKDEED-----EEDKEDEEDVPQAKDEL	280
Db	380	GGDEDDDDLEEDDEKADKADSKDSDEKQ-----HDEL	421

Search completed: October 21, 2001, 03:07:45
Job time: 340 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:08:25 ; Search time 28.98 Seconds
(without alignments)
330.971 Million cell updates/sec

Title: US-09-807-148-9
Perfect score: 1553
Sequence: 1 GPGTKVHVIFNYKKNVLI.....EEDKEDEEDVPGQAKDEL 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553	100.0	417	1 CRTC_HUMAN	P27797 homo sapien
2	1487.5	95.8	416	1 CRTC_MOUSE	P14211 mus musculu
3	1476.5	95.1	416	1 CRTC_RAT	P18418 rattus norv
4	1457.5	93.9	418	1 CRTC_RABIT	P15253 oryctolagus
5	1442	92.9	400	1 CRT1_BOVIN	P52193 bos taurus
6	1432	92.2	421	1 CRT2_BOVIN	P42918 bos taurus
7	1031.5	66.4	406	1 CRTC_DROME	P29413 drosophila
8	991.5	63.8	395	1 CRTC_CAEEL	P27798 caenorhabdi
9	944.5	60.8	388	1 RAL1_ONCVO	P11012 onchocerca
10	856.5	55.2	415	1 CRTC_RICCO	P93508 ricinus com
11	837.5	53.9	420	1 CRTC_CHLRE	Q9std3 chlamydomon
12	837.5	53.9	420	1 CRTC_MAIZE	Q9sp22 zea mays (m
13	824	53.1	424	1 CRT2_ARATH	Q38858 arabidopsis
14	821.5	52.9	421	1 CRTC_PRUAR	Q9xf98 prunus arne
15	818.5	52.7	416	1 CRTC_NICPL	Q0401 nicotiana p
16	817.5	52.6	393	1 CRTC_SCHMA	Q06814 schistosoma
17	814.5	52.4	416	1 CRTC_BETVO	O81919 beta vulgar
18	810.5	52.2	425	1 CRT1_ARATH	O04151 arabidopsis
19	795	51.2	401	1 CRTC_EUGER	Q9znv3 euglena gra
20	790.5	50.9	416	1 CRTC_BERST	Q9zpp1 berberis st
21	779	50.2	424	1 CRTC_ORYSA	Q9slv8 oryza sativ
22	764	49.2	424	1 CRTC_DICDI	Q23858 dictyostell
23	699.5	45.0	424	1 CRT3_ARATH	O04153 arabidopsis
24	633	40.8	591	1 CALX_MOUSE	P35564 mus musculu
25	629	40.5	591	1 CALX_RAT	P35565 rattus norv
26	628	40.4	592	1 CALX_HUMAN	P27824 homo sapien
27	624	40.2	593	1 CALX_CANFA	P24843 canis famli
28	577.5	37.2	610	1 CALG_HUMAN	O14567 homo sapien
29	534	35.7	611	1 CALG_MOUSE	P52194 mus musculu
30	531	34.2	546	1 CALX_SOYBN	Q39817 glycine max
31	501.5	32.3	530	1 CAX1_ARATH	P29402 arabidopsis
32	493	31.7	551	1 CALX_PEA	O82709 pisum sativ
33	479.5	30.9	528	1 CAX2_ARATH	Q38798 arabidopsis

ALIGNMENTS

RESULT 1

ID	CRTC_HUMAN	STANDARD;	PRT;	417 AA.
AC	P27797;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (52 KDA			
DE	RIBONUCLEOPROTEIN AUTOANTIGEN RO/SS-A).			
GN	CALR OR CRTC.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=92013129; PubMed=1919005;			
RA	Rokeach L.A., Haselby J.A., Meilof J.F., Smeenk R.J., Unnasch T.R.,			
RA	Greene B.M., Hoch S.O.;			
RT	"Characterization of the autoantigen calreticulin.";			
RL	J. Immunol. 147:3031-3039(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=90237213; PubMed=2332496;			
RA	McCaulliffe D.P., Lux F.A., Liew T.S., Sanz I., Hanke J., Newkirk M.M.,			
RA	Bachinski L.L., Itoh Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,			
RA	Capra J.D.;			
RT	"Molecular cloning, expression, and chromosome 19 localization of a			
RT	human RO/SS-A autoantigen.";			
RL	J. Clin. Invest. 85:1379-1391(1990).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=92129342; PubMed=1733953;			
RA	McCaulliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;			
RT	"The 5'-flanking region of the human calreticulin gene shares			
RT	homology with the human GRP78, GRP94, and protein disulfide isomerase			
RT	promoters.";			
RL	J. Biol. Chem. 267:2557-2562(1992).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RP	Lamerdin J., McCreedy P., Stilwagen S., Ramirez M., Carrano A.;			
RP	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE OF 18-36.			
RP	MEDLINE=92002034; PubMed=1911778;			
RA	Rojlani M.V., Finlay B.B., Gray V., Dedhar S.;			
RT	"In vitro interaction of a polypeptide homologous to human RO/SS-A			
RT	antigen (calreticulin) with a highly conserved amino acid sequence in			
RT	the cytoplasmic domain of integrin alpha subunits.";			
RL	Biochemistry 30:9859-9866(1991).			
RN	[6]			
RP	SEQUENCE OF 18-32.			
RP	MEDLINE=90380058; PubMed=2400400;			
RA	Krause K.H., Simmerman H.K.B., Jones L.R., Campbell K.P.;			
RT	"Sequence similarity of calreticulin with a Ca2(+)-binding protein			
RT	that co-purifies with an Ins(1,4,5)P3-sensitive Ca2+ store in HL-60			

P36581 schizosacch
P34652 caenorhabdi
Q39994 helianthus
P27825 saccharomyc
P08199 mesocricetu
P91753 lytechinus
P22620 plasmodium
P09405 mus musculu
P15771 gallus gall
P13383 rattus norv
P23746 plasmodium
P06180 xenopus lae

34 472.5 30.4 1 CALX_SCHPO
35 472.5 30.4 1 CALX_CAEEL
36 451.5 29.1 540 1 CALX_HELTO
37 299.5 19.3 502 1 CALX_YEAST
38 184.5 11.9 713 1 NUCCL_MESAU
39 184 11.8 411 1 MP62_LVTPI
40 183.5 11.8 743 1 ABRA_PLAFC
41 177.5 11.4 706 1 NUCCL_MOUSE
42 176 11.3 694 1 NUCCL_CHICK
43 175.5 11.3 712 1 NUCCL_RAT
44 173.5 11.2 321 1 ABRA_PLAFA
45 173.5 11.2 589 1 HIBN_XENLA

CC LOW AFFINITY CALCIUM-BINDING SITES.

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.

CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: X14926; CAA33053.1; -

CC EMBL: M92988; AAA37569.1; -

CC PIR: S06763; S06763.

CC PIR: JCI1444; JCI1444.

CC SWISS-2DPAGE: P14211; MOUSE.

CC MGI: 88252; Calr.

CC InterPro: IPR000886; -

CC InterPro: IPR001580; -

CC Pfam: PF00262; calreticulin; 1.

CC PRINTS: PR00626; CALRETICULIN.

CC PROSITE: PS00014; ER_TARGET; 1.

CC PROSITE: PS00803; CALRETICULIN_1; 1.

CC PROSITE: PS00804; CALRETICULIN_2; 1.

CC PROSITE: PS00805; CALRETICULIN_REPEAT; 3.

CC Endoplasmic reticulum; Calcium-binding; Repeat; Signal.

CC SIGNAL 1 17

CC CHAIN 18 416 CALRETICULIN.

CC DOMAIN 18 197 N-DOMAIN.

CC DOMAIN 198 308 P-DOMAIN.

CC DOMAIN 309 416 C-DOMAIN.

CC DOMAIN 191 255 4 X APPROXIMATE REPEATS.

CC REPEAT 191 202 1-1.

CC REPEAT 210 221 1-2.

CC REPEAT 227 238 1-3.

CC REPEAT 244 255 1-4.

CC REPEAT 259 297 3 X APPROXIMATE REPEATS.

CC REPEAT 259 269 2-1.

CC REPEAT 273 283 2-2.

CC REPEAT 287 297 2-3.

CC DOMAIN 351 407 ASP/GLU/LYS-RICH.

CC DISULFID 137 163 BY SIMILARITY.

CC SITE 413 416 PREVENT SECRETION FROM ER.

CC SEQUENCE 416 AA: 47994 MW: 24C03B00913408D8 CRC64;

Query Match 95.88; Score 1487.5; DB 1; Length 416;

Best Local Similarity 95.48; Pred. No. 1.9e-79;

Matches 267; Conservative 9; Mismatches 3; Indels 1; Gaps 1;

Qy 1 GPGTKKVVHIFNYKGNVINKIRCKDDPTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 60

Db 138 GPGTKKVVHIFNYKGNVINKIRCKDDPTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 197

Qy 61 DDWDFLPKKTKDPAKPEDWDERAKIDDDPTDSKPEDWDERAKIHPPDPAKPEDWDEEM 120

Db 198 DDWDFLPKKTKDPAKPEDWDERAKIDDDPTDSKPEDWDERAKIHPPDPAKPEDWDEEM 257

Qy 121 DGEWEPVIONPEYKGEWKPKQIDNPQYKGTWIIHPIIDNPYSPDPSIYAYDNFGVLGLD 180

Db 258 DGEWEPVIONPEYKGEWKPKQIDNPQYKGTWIIHPIIDNPYSPDPSIYAYDNFGVLGLD 317

Qy 181 LWQKSGTIFNFLTITNDEAYAEFGNETGWTKAAEKOMKDKODEORLKEEEDKKRK 240

Db 318 LWQKSGTIFNFLTITNDEAYAEFGNETGWTKAAEKOMKDKODEORLKEEEDKKRK 377

Qy 241 EEEAEKDEDDKDE 280

Db 378 EEEAEKDEDDKDE 416

RESULT . 3

CRTC_RAT

ID CRTC_RAT STANDARD; PRT; 416 AA.

AC P18418; P10452;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE CALRETICULIN PRECURSOR (CRP55) (CALREGULIN) (HACBP) (ERP60) (CALBP)

DE (CALCIUM-BINDING PROTEIN 3) (CABP3).

GN CALR.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain cortex;

RX MEDLINE=90370496; PubMed=2395661;

RA Murthy K.K., Banville D., Srikanth C.B., Carrier F., Bell A.,

RA Holmes C., Patel Y.C.;

RA "Structural homology between the rat calreticulin gene product and

RT the Onchocerca volvulus antigen Ral-1.";

RL Nucleic Acids Res. 18:4933-4933(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RX MEDLINE=93202172; PubMed=8453984;

RA Nakamura M., Moriya M., Baba T., Michikawa Y., Yamanobe T., Arai K.,

RA Okinaga S., Kobayashi T.;

RA "An endoplasmic reticulum protein, calreticulin, is transported into

RT the acrosome of rat sperm.";

RL Exp. Cell Res. 205:101-110(1993).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;

RX MEDLINE=95181573; PubMed=7876339;

RA Soennichsen B., Fueflektug J., van Nguyen P., Diekmann W.,

RA Robinson D.G., Mieskes G.;

RA "Retention and retrieval: both mechanisms cooperate to maintain

RT calreticulin in the endoplasmic reticulum.";

RL J. Cell Sci. 107:2705-2717(1994).

RN [4]

RP SEQUENCE OF 270-358 FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RA Lone Y.C., Bailly A., Latruffe N.;

RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE OF 18-29.

RX MEDLINE=91054414; PubMed=2241926;

RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,

RA MacLennan D.H., Meldolesi J., Pozzan T.;

RA "Calreticulin is a candidate for a calsequestrin-like function in

RT Ca2(+)-storage compartments (calciosomes) of liver and brain.";

RL Biochem. J. 271:473-480(1990).

RN [6]

RP SEQUENCE OF 18-32.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;

RX MEDLINE=92360010; PubMed=1497655;

RA Nakamura M., Michikawa Y., Baba T., Okinaga S., Arai K.;

RA "Calreticulin is present in the acrosome of spermatids of rat

RT testis.";

RL Biochem. Biophys. Res. Commun. 186:668-673(1992).

RN [7]

RP SEQUENCE OF 18-32.

RC STRAIN=LEC; TISSUE=Liver;

RX MEDLINE=94072621; PubMed=8251535;

RA Yokoi T., Nagayama S., Kajiwara R., Kawaguchi Y., Horiuchi R.,

RA Kamataki T.;

RA "Identification of protein disulfide isomerase and calreticulin as

RT autoimmature antigens in LEC strain of rats.";

RL Biochim. Biophys. Acta 1158:339-344(1993).

CC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND

CC LOW AFFINITY CALCIUM-BINDING SITES.

RA Zheng X.H., Zhong F.N., Zhong W.F., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2185-2195(2000).
 RN [3]
 RN
 RN SEQUENCE OF 91-124 AND 182-220.
 RP.
 RX MEDLINE=90307981; PubMed=2365822;
 RA McCaulliffe D.P., Zappi E., Lieu T.S., Michalak M., Sontheimer R.D.,
 RA Capra J.D.;
 RA "A human Ro/SS-A autoantigen is the homologue of calreticulin and is
 RT highly homologous with onchocercal RAL-1 antigen and an alypsia
 RT 'memory molecule'";
 RT J. Clin. Invest. 86:332-335(1990).
 RL CC
 RL CC -I- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
 CC LOW AFFINITY CALCIUM-BINDING SITES.
 CC -I- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -I- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X64461; CA445791.1; -;
 DR EMBL; AE003683; AAF54416.1; -;
 DR PIR; A37158; A37158.
 DR FlyBase; FBgn0005585; Crc:
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PROSITE; PS00014; ER_TARGET; 1.
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 406 CALRETICULIN.
 FT CONFLICT 107 107 G -> A (IN REF. 3).
 FT CONFLICT 184 184 V -> L (IN REF. 3).
 SQ SEQUENCE 406 AA; 46808 MW; 65D72C69D08EC427 CRC64;

Query Match	66.4%;	Score 1031.5;	DB 1;	Length 406;
Best Local Similarity	67.3%;	Pred. No. 3.5e-53;		
Matches 181;	Conservative 36;	Mismatches 51;	Indels 1;	Gaps 1;
QY	1	GPGTKKVHVITFNKYGKNVLINKOIRCKDDFETHLYTLIVRPDNTFYEVKIQNSQVESGSL	60	
		: : : : : : : : :		
Db	138	GPGTKKVHVITSYKGKNHLISKIOIRCKDDVTYFTYTLIVRPDNTFYELIDNEKVESGNLE	197	
		: : : : : : : :		
QY	61	DWDNDFLPPKKIKDPDASKPEDWDERAKIDDPTDSKPEDWDKPEHIPDPDAKKPDEWDEEM	120	
		: : : : : : : :		
Db	198	DWDNDFLAPKKIKDPTATPKPEDWDORATIPDPDDKKPEDWDKPEHIPDPDATKPDWDDEM	257	
		: : : : : : : :		
QY	121	DGEWEPPVIONPCKYGBWKPRQINDPYDKTWTHTPETDNPEYSPDPISYIAYDNFGVLGLD	180	
		: : : : : : : :		
Db	258	DGEWEPPMINPEFKGEWGPKQLDNPNYKGAWHEPIANPEYVPDDKLYLRKETCTJTGFD	317	
		: : : : : : : :		
QY	181	LWQVKSGSTIFDNFLINDEAYAEFFNGNETWGVTKAAECHKMDQEQRULKEEBEDKRK	240	
		: : : : : : : :		
Db	318	LWQVKSGSTIFDNVLITDDVELAATAAAAEVN-TQAGEKKMKAEQDEVORKDDEEAKKAS	376	
		: : : : : : : :		
QY	241	EBEAEADKDEDKDDEDEEDEEKDEE	269	
		: : : : : : : : : : : : : : : : : : : :		
Db	377	DKDEDDEDDDEERKDDSKODKDOSEHDE	405	

RESULT	8
CRTC_CAEEL	

ID	CRTC_CABEEL	STANDARD;	PRT;	395 AA.
DT	.P27798;			
DT	01-AUG-1992	(Rel. 23, Created)		
DT	01-AUG-1992	(Rel. 23, Last sequence update)		
DT	01-FEB-1994	(Rel. 28, Last annotation update)		
DE	CALRETICULIN PRECURSOR.			
GN	CRT-1.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BRISTOL N2;			
RX	MEDLINE=92329978; PubMed=1627827;			
RA	Smith M.J.;			
RT	"A C. elegans gene encodes a protein homologous to mammalian			
RT	calreticulin".;			
RL	DNA Seq. 2:235-240(1992).			
CC	-1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND			
CC	LOW AFFINITY CALCIUM-BINDING SITES.			
CC	-1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.			
CC	-1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.			
CC	-----			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X59589; CAA42159.1; -			
DR	PIR; S25851; S25851			
DR	InterPro; IPR000886; -			
DR	InterPro; IPR001580; -			
DR	Pfam; PF00262; calreticulin; 1.			
DR	PRINTS; PR00626; CALRETICULIN.			
DR	PROSITE; PS00014; ER_TARGET; 1.			
DR	PROSITE; PS00803; CALRETICULIN_1; 1.			
DR	PROSITE; PS00804; CALRETICULIN_2; 1.			
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.			
KW	Endoplasmic reticulum; Calcium-binding; Repeat; Signal.			
FT	SIGNAL 1	POTENTIAL.		
FT	CHAIN ?	395		
FT	CHAIN ?	N-DOMAIN.		
FT	DOMAIN ?	192		
FT	DOMAIN	193 301		
FT	DOMAIN	302 395		
FT	DOMAIN	186 250		
FT	REPEAT	186 197		
FT	REPEAT	205 216		
FT	REPEAT	222 233		
FT	REPEAT	239 250		
FT	DOMAIN	254 292		
FT	REPEAT	254 264		
FT	REPEAT	268 278		
FT	REPEAT	282 290		
FT	DOMAIN	332 390		
FT	DISULFID	133 158		
FT	SITE	392 395		
SQ	SEQUENCE 395 AA; 45616 MW; 35CA7D2EC1D56B03 CRC64;			

Query Match	63.8%	Score	991.5;	DB	1;	Length	395;
Best Local Similarity	66.0%;	Pred.	No. 7e-51;				
Matches	173;	Conservative	35;	Mismatches	53;	Indels	1;
Gaps	1;						
QY	1	GPGTKKVHVIFNFKGNVLINKOIRCDDEETHLYLTILIVPNDTYEYKIDNSQVESGSL	60				
		: : : :					
Dd	134	GP-TRRVHVITNTYKGKNLKJKEITCKSDSETHLYTLILNSDNTYEYKIDGESAQTSGL	192				
		:					
QY	61	DDWDFLPKKIKDPDASKPEDWDERAKIDDPDTDSKPEDWDKEPHIPDPADKA	120				
		: : : : : : : : : : : : :					

2006

Db 193 EDWLLPAKKIKDPDAKKEDWDEREYIDDAEDAKPEDWEKPEHIPPDPDAKKPEDWDEM 252
QY 121 DGEWEPVTONPEYKGEWPRQIDNDYKGTWTHPEIDNPEYSPDPSIYAYDNFGVLGLD 180
Db 253 DGEWEPVTONPEYKGEWPKQIKNPAYKGLWHPHENPEYTPDDELYSYESWGAIGFD 312
QY 181 LMOVKSCTIFDNLITNDAYAEFGNETWGTAAEKQKQDEORLKEEEDKKRK 240
Db 313 LMOVKSCTIFDNLITNDAYAEFGNETWGTAAEKQKQDEORLKEEEDKKRK 240
QY 241 EEEEEEDKEDDEDEE 262
Db 373 EEKAKDDDEEEKEEGHDE 394

RESULT 9

RAL1_ONCVO STANDARD; PRT; 388 AA.
ID RAL1_ONCVO STANDARD; PRT; 388 AA.
AC P11012;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RAL-1 PROTEIN PRECURSOR (41 KDA LARVAL ANTIGEN).
GN RAL1.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94341871; PubMed=7520419;
RA Rokeach L.A., Zimmerman P.A., Unnasch T.R.;
RT "Epitopes of the Onchocerca volvulus RAL1 antigen, a member of the
RT calreticulin family of proteins, recognized by sera from patients
RT with onchocerciasis".
RL Infect. Immun. 62:3696-3704(1994).
RN [2]

RP SEQUENCE OF 53-388 FROM N.A.
RX MEDLINE=88273584; PubMed=2455736;
RA Unnasch T.R., Gallin M.X., Soboslay P.T., Ertmann K.D., Greene B.M.;
RT "Isolation and characterization of expression cDNA clones encoding
RT antigens of Onchocerca volvulus infective larvae".
RL J. Clin. Invest. 82:262-269(1988).
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20565; AA59056.1; -
CC PIR; A32507; A32507.
CC
CC InterPro; IPR001580; -
CC Pfam; PF00262; calreticulin; 1.
CC PRINTS; PR00626; CALRETICULIN.
CC PROSITE; PS00803; CALRETICULIN_1; 1.
CC PROSITE; PS00804; CALRETICULIN_2; 1.
CC PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
CC Calcium-binding; Repeat; Antigen; Signal.
CC SIGNAL 1 17
CC CHAIN 18 388 RAL-1 PROTEIN.
CC FT DOMAIN 189 253 4 X APPROXIMATE REPEATS.
CC FT REPEAT 189 200 1-1.
CC FT REPEAT 208 219 1-2.
CC FT REPEAT 225 236 1-3.
CC FT REPEAT 242 253 1-4.
CC FT DOMAIN 257 295 3 X APPROXIMATE REPEATS.
CC FT REPEAT 257 267 2-1.
CC FT REPEAT 271 281 2-2.
CC FT REPEAT 285 295 2-3.

FT DOMAIN 353 388 ARG/LYS-RICH (BASIC).
FT DISULFID 135 161 BY SIMILARITY.
SQ SEQUENCE 388 AA; 45298 MW; 9537F298A2D31CD6 CRC64;

Query Match 60.8%; Score 944.5; DB 1; Length 388;
Best Local Similarity 65.4%; Pred. No. 3.5e-48;
Matches 166; Conservative 32; Mismatches 49; Indels 7; Gaps 2;

QY 1 GPGTKVHVIFENYKGNVLINKDIRCKDDVFTHLVTLIVRPONTYEVKIDNSOVESGSL 60
Db 136 GPGTKVHVIFHYKDNHMKKDIRCKDDVFTHLVTLIVNSDNTYEVQIDGEAESGELE 195
QY 61 DDWDFLPPKKIKDPDAKSKPEDDERAKIDDPDTSKPEDWPKPEHIPPDPDAKKPEDWDEM 120
Db 196 ADWDFLPPKKIKDPDAKSKPEDDEREFIDDEDKPEDWPKPEHIPPDPDAKKPEDWDEM 255
QY 121 DGEWEPVTONPEYKGEWPRQIDNDYKGTWTHPEIDNPEYSPDPSIYAYDNFGVLGLD 180
Db 256 DGEWEPVTONPEYKGEWPKQIKNPAYKGLWHPHENPEYTPDDELYSYESWGAIGFD 312
QY 181 LMOVKSCTIFDNLITNDAYAEFGNETWGTAAEKQKQDEORLKEEEDKKRK 240
Db 313 LMOVKSCTIFDNLITNDAYAEFGNETWGTAAEKQKQDEORLKEEEDKKRK 240
QY 237 --KKRKEEEAEEDK 248
Db 375 KMKRKRANRKKKK 388

RESULT 10

CRTC_RICCO
ID CRTC_RICCO STANDARD; PRT; 415 AA.
AC P93508;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97435975; PubMed=9290642;
RA Coughlan S.J., Hastings C., Winfrey R. Jr.;
RT "Cloning and characterization of the calreticulin gene from Ricinus
RT communis L".
RL Plant Mol. Biol. 34:897-911(1997).
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES.
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U74631; AA871420.1; -
CC DR EMBL; U74630; AA871419.1; -
CC Mendel; 10452; Ricco:1166;10452.
CC InterPro; IPR000886; -
CC InterPro; IPR001580; -
CC Pfam; PF00262; calreticulin; 1.
CC PRINTS; PR00626; CALRETICULIN.
CC PROSITE; PS00014; ER_TARGET; 1.
CC PROSITE; PS00803; CALRETICULIN_1; 1.
CC PROSITE; PS00804; CALRETICULIN_2; 1.

DR	PROSITE; PS00803; CALRETICULIN_1; 1.
DR	PROSITE; PS00804; CALRETICULIN_2; 1.
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 1.
KW	Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT	SIGNAL 1 18 POTENTIAL.
FT	CHAIN 19 420 CALRETICULIN.
FT	STATE 417 420 PREVENT SECRETION FROM ER (POTENTIAL).
SQ	SEQUENCE 420 AA, 47327 MW; DD3BAFFBF61C9B CRC64;

	Query Match	Score 837.5;	DB 1;	Length 420;
	Best Local Similarity	53.9%;		
	Matches 154;	54.0%;	Pred. No. 5.6e-42;	
	Conservative	43;	Mismatches 65;	Indels 23;
				Gaps 4;

[illegible]

```

RESULT 12
CRTC_MAIZE ID CRTC_MAIZE STANDARD; PRT; 420 AA.
AC Q9SP22;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALRETICULIN PRECURSOR.
GN CRT.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
ON NCBI_TaxID=4577;
RX [1]
RP SEQUENCE FROM N.A.
RT WYatt W.E., Tsou P.-L., Robertson D.;
RT "Effects of altered expression of the calcium-binding protein
RT calreticulin in Arabidopsis thaliana.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RC -1- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF190454; AAF01470.1; -.
DR InterPro: IPR000886; -.
DR IntraPro: IPR001580; -.
DR Pfam: PF00262; Calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.

```

```
DR PROSITE: PS00014; ER_TARGET: 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; FALSE_NEG.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 420 CALRETICULIN.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SITE 417 420 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 420 AA; 47939 MW; E73B7F43E7494735 CRC64;

Query Match 53.9%; Score 837.5; DB 1; Length 420;
Best Local Similarity 52.9%; Pred. No. 5.6e-42;
Matches 147; Conservative 49; Mismatches 77; Indels 5; Gaps 2;

QY 4 TTKVHVIFNYKGNVNLINKDIRCKDDEFTLTVLPDNTVEYKIDNSQVSGSLEDOW 63
Db 147 TTKVHVIFNYKGNVNLINKDIRCKDDEFTLTVLPDNTVEYKIDNSQVSGSLEDOW 63
QY 64 DELPPKKIKDPDPAKPEDWDERAKIDDPDTSKPEDWK-PEHIPDPDAKPPEDWDEMDG 122
Db 207 DILPPKKIKDPDPAKPEDWDERAKIDDPDTSKPEDWK-PEHIPDPDAKPPEDWDEMDG 122
QY 123 EWEPPVQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTYAYDNFVGLGLDW 182
Db 267 EWTAPTIPNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTYAYDNFVGLGLDW 182
QY 183 QVKSQGTIFDNFLITNDDEAYAEFGNETGWTKAAEKQKDKDEEORLKEFEEDKKRKEE 242
Db 327 QVKSQGTIFDNFLITNDDEAYAEFGNETGWTKAAEKQKDKDEEORLKEFEEDKKRKEE 242
QY 243 EEADEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
Db 383 DEEDDLEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280

RESULT 13
CRT2_ARATH
ID CRT2_ARATH STANDARD; PRT; 424 AA.
AC Q3858; 080486; 004152;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE CALRETICULIN 2 PRECURSOR.
GN CRT2 OR CRTL OR t12M4.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaya V.S., Schwartz J.R., Toriumi M., Yu G., Oji O., Kwan A.,
RA Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B.,
RA Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,
RA Shinn P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,
RA Theologis A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-174 FROM N.A.
RX MEDLINE=97303616; PubMed=9159940;
RA Nelson D.E., Glaunsinger B., Bohnert H.J.;
RT "Abundant accumulation of the calcium-binding molecular chaperone
RT calreticulin in specific floral tissues of Arabidopsis thaliana.";
RL Plant Physiol. 114:29-37(1997).
RN [3]
RP SEQUENCE OF 16-424 FROM N.A.
RC STRAIN=CV. LANDSBERG RECTA; TISSUE=Flower;
RA Benedetti C.E., Turner J.G.;
RT "Nucleotide sequence of an Arabidopsis thaliana cDNA encoding a
RT protein homologous to plant and animal calreticulins.";
```

```
RL (In) Plant Gene Register PGR95-047.
CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND
CC LOW AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY,
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AC003114; AAC24083.1; ALT SEQ.
DR EMBL; U66344; AAC49696.1; ALT_INIT.
DR EMBL; U27698; AAC80652.1; -.
DR HSSP; P00288; ARXN.
DR SWISS-2DPAGE; Q38858; ARATH.
DR Mendel; 31147; Arath;1166;31147.
DR Mendel; 16506; Arath;1166;16506.
DR InterPro; IPR000886; -.
DR InterPro; IPR001560; -.
DR Pfam; PF00262; calreticulin; 1.
DR PROSITE; PS00636; CALRETICULIN.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 2.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein;
KW Multigene family.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 424 CALRETICULIN 2.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SITE 421 424 PREVENT SECRETION FROM ER (POTENTIAL).
FT CONFLICT 16 19 LVAI -> NSAR (IN REF. 3).
FT CONFLICT 155 155 P -> E (IN REF. 1).
FT CONFLICT 236 236 P -> T (IN REF. 3).
FT CONFLICT 408 408 V -> E (IN REF. 3).
SQ SEQUENCE 424 AA; 48084 MW; 514385EBAE810DD7 CRC64;

Query Match 53.1%; Score 824; DB 1; Length 424;
Best Local Similarity 54.0%; Pred. No. 3.4e-41;
Matches 155; Conservative 36; Mismatches 86; Indels 10; Gaps 5;

QY 1 GPGTKKVVHIFNYKGNVNLINKDIRCKDDEFTLTVLPDNTVEYKIDNSQVSGSLE 60
Db 141 GYSTKKVHAILTYNGANHLIKKDVPCETDQLTHVYFILRPDATYSLIDNVEKQTGSLY 200
QY 61 DDMDFLPPKKIKDPDPAKPEDWDERAKIDDPDTSKPEDWK-PEHIPDPDAKPPEDWDEE 119
Db 201 SDWLLPPKKIKDPDPAKPEDWDEYISDPDKKPDGVDIIPKEIPDPTDSKPPEDWDE 260
QY 120 MGEWEPPVQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTYAYDNFVGLGL 179
Db 261 EDGEWTAPTIPNPYKGEWKPRQIDNPDKYKGTWHPIDNPDKPDPELYVFPKLYVGL 320
QY 180 DLWQVKSQGTIFDNFLITNDDEAYAEFGNETGWTKAAEKQKDKDEEORLKEFEEDKKR 239
Db 321 ELWQVKSQGTIFDNFLITNDDEAYAEFGNETGWTKAAEKQKDKDEEORLKEFEEDKKR 239
QY 240 KKEEAEEDK-EDDE---DKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
Db 378 PAESDAEDEPEDEGGDDSDSESKEAETKSVDSSEETSEKDATAHDEL 424

RESULT 14
CRTC_PUAR
ID CRTC_PUAR STANDARD; PRT; 421 AA.
AC Q9XF98;
DT 01-OCT-2000 (Rel. 40, Created)
```

[illegible]

Db 375 EEEESKAAPADSDAEEDDDDDDDDDDDKSEKDD----EAHDEL 416

Search completed: October 21, 2001, 03:08:26
Job time: 276 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:10:09 ; Search time 92.24 Seconds
(without alignments)
401.620 Million cell updates/sec

Title: US-09-807-148-9
Perfect score: 1553
Sequence: 1 GPGTKKVVHVFNYKGNVLI.....EDKKEDEEDVPGQAKDEL 280

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTRMBL16.*
- 2: sp_archaea.*
- 3: sp_bacteria.*
- 4: sp_fungi.*
- 5: sp_human.*
- 6: sp_invertebrate.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_unclassified.*
- 13: sp_vertebrate.*
- 14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1280.5	82.5	411	13 Q91710	Q91710 xenopus lae
2	1241.5	79.9	419	13 Q98984	Q98984 rana rugosa
3	1184	76.2	318	13 Q9PTX7	Q9PTX7 lampetra re
4	1164	75.0	405	5 Q26268	Q26268 aplysia cal
5	1136	73.1	417	13 Q9PUC1	Q9PUC1 brachydanio
6	1120	72.1	343	13 Q91711	Q91711 xenopus lae
7	1095.5	70.5	321	13 Q9U5G0	Q9U5G0 eptatretus
8	1093.5	70.4	410	5 Q16893	Q16893 amblyomma a
9	1049.5	67.6	421	5 Q9U6S0	Q9U6S0 strongyloce
10	1033	66.5	403	5 Q76961	Q76961 necator ame
11	1026.5	66.1	406	5 Q9U916	Q9U916 drosophila
12	997.5	64.2	375	5 Q18478	Q18478 litomosoide
13	966.5	62.2	387	5 Q97372	Q97372 dirofilaria
14	844	54.3	421	10 Q43712	Q43712 zea mays (m
15	843.5	54.3	321	10 Q41799	Q41799 zea mays (m
16	836.5	53.9	427	10 Q9FV2	Q9FV2 pinus taeda
17	833	53.6	396	5 Q45034	Q45034 schistosoma
18	817.5	52.6	389	10 Q40567	Q40567 nicotiana t
19	817	52.6	350	5 Q26514	Q26514 schistosoma

ALIGNMENTS

RESULT 1

```

ID Q91710 PRELIMINARY; PRT; 411 AA.
AC Q91710;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE CALRETICULIN PRECURSOR (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC T'SSUE=BRAIN;
RX MEDLINE=93074997; PubMed=1445218;
RA Treves S., Zorzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous
   system.";
RL Biochem. J. 287:579-581(1992).
DR EMBL: X67597; CAA47866.1; -
DR InterPro: IPR000886; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
DR Signal.
KW NON_TER
FT SIGNAL
FT CHAIN
SQ SEQUENCE 411 AA; 48344 MW; 891DA66E00EBEFA CRC64;

```

Query Match 82.5%; Score 1280.5; DB 13; Length 411;
Best Local Similarity 80.1%; Pred. No. 1.7e-84;
Matches 225; Conservative 31; Mismatches 22; Indels 3; Gaps 2;

QY 1 GPGTKKVVHVFNYKGNVLI.....EDKKEDEEDVPGQAKDEL 280

```

Db 133 GPTTKVHVIFQYKKNLQINDIRCKDDSFTHLYTLIVRPONTYEVKIDNSKVESGSLE 192
QY 61 DDWDFLPPKKIKDPDASKPEDDERAKIDDPDTSKPEDWDPKPEHIPDPDAKPEDWDEEM 120
Db 193 DDWDFLPPKKIKDPDPAKPEDWDERPKIDDPEDKPEDWKEPEHIPDPDAKPEDWDEEM 252
QY 121 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFGVLGLD 180
Db 253 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFGVLGLD 312
QY 181 LMQVKSFTFDNFLTNDAYAEFGNETGWVTKAAEKQMKQKQDEEQRLEKEEEDKKRK 240
Db 313 LMQVKSFTFDNFLTNDKHAEEYGNETGWVTKAAEKQMKQKQDEEQRLEKEEEDKKRK 372
QY 241 EEEAEEDKEDDEDKDEDEEDKDEEEDVPGQAKDEL 280
Db 373 EEEAEEDKEDDEDKDEDEEDKDEEEDVPGQAKDEL 411

RESULT 2
Q98984
ID Q98984 PRELIMINARY; PRT; 419 AA.
AC Q98984
DT 01-FEB-1997 (TremBLrel. 02, Created)
DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE CALRETICULIN.
OS Rana rugosa (Frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=84110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96234004; PubMed=8654561;
RA Yamamoto S., Nakamura M.;
RT "Calnexin: its molecular cloning and expression in the liver of the
RT frog, Rana rugosa.";
RL FEBS Lett. 387:27-32(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Yamamoto S.;
RT "Strong expression of the calreticulin gene in the liver of Rana
RT rugosa tadpoles.";
RL J. Exp. Zool. 0:0-0(1996).
DR EMBL; D78589; BAA11425.1;
DR InterPro; IPR000886;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866;
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 419 AA; 48658 MW; 2C857036769673BF CRC64;

Query Match 79.9%; Score 1241.5; DB 13; Length 419;
Best Local Similarity 76.9%; Pred. No. 1.le-81;
Matches 216; Conservative 34; Mismatches 30; Indels 1; Gaps 1;

QY 1 GPGTKKVHVIFNYKGNVLINKDIRCKDEFFTHLYTLIVRPONTYEVKIDNSQVESGSLE 60
Db 139 GPGTKKVHVIFNYKGNVLINKDIRSKADYSHLYTLIVRPONTYEVKIDNSKVESGNLE 198
QY 61 DDWDFLPPKKIKDPDASKPEDDERAKIDDPDTSKPEDWDPKPEHIPDPDAKPEDWDEEM 120
Db 199 DDWDFLPPKKIKDPDPAKPEDWDERPKIDDPEDKPEDWKEPEHIPDPDAKPEDWDEEM 258
QY 121 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFGVLGLD 180
Db 259 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFGVLGLD 318

```

```

QY 181 LMQVKSFTFDNFLTNDAYAEFGNETGWVTKAAEKQMKQKQDEEQRLEKEEEDKKRK 240
Db 319 LMQVKSFTFDNFLTNDKFAEEHATKTGWVTKGEKKMKQKQDEEQRLEKEEEDKKRK 378
QY 241 EEEAEADKEDDEDKDEDEEDKDEEEDVPGQAKDEL 280
Db 379 EEEAEADKEDDEDKDEDEEDKDEEEDVPGQAKDEL 419

RESULT 3
Q9PTX7
ID Q9PTX7 PRELIMINARY; PRT; 318 AA.
AC Q9PTX7;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Lampetra reissneri.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7753;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20063780; PubMed=10594174;
RA Kuraku S., Hoshiyama D., Katoh K., Suga H., Miyata T.;
RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded
RT genes.";
RL J. Mol. Evol. 49:729-735(1999).
DR EMBL; AB025328; BAA88481.1;
DR InterPro; IPR000886;
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR PRODOM; PD001866;
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
DR NON_TER 1
FT SEQUENCE 318 AA; 36997 MW; C88102EA1CAC1506 CRC64;

Query Match 76.2%; Score 1184; DB 13; Length 318;
Best Local Similarity 73.1%; Pred. No. 1.e-77;
Matches 207; Conservative 36; Mismatches 36; Indels 4; Gaps 2;

QY 1 GPGTKKVHVIFNYKGNVLINKDIRCKDEFFTHLYTLIVRPONTYEVKIDNSQVESGSLE 60
Db 37 GPGTKKVHVIFNYKGNVLINKDIRCKDEFFTHLYTLIVRPONTYEVKIDNSKVESGSLE 96
QY 61 DDWDFLPPKKIKDPDASKPEDDERAKIDDPDTSKPEDWDPKPEHIPDPDAKPEDWDEEM 120
Db 97 DDWDFLPPKKIKDPDPAKPEDWDERPKIDDPEDKPEDWKEPEHIPDPDAKPEDWDEEM 156
QY 121 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFGVLGLD 180
Db 157 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFGVLGLD 216
QY 181 LMQVKSFTFDNFLTNDAYAEFGNETGWVTKAAEKQMKQKQDEEQRLEKEEEDKKRK 240
Db 217 LMQVKSFTFDNFLTNDKFAEEHATKTGWVTKGEKKMKQKQDEEQRLEKEEEDKKRK 276
QY 241 EEEAEADKEDDEDKDEDEEDKDEEEDVPGQAKDEL 280
Db 277 EEEAEADKEDDEDKDEDEEDKDEEEDVPGQAKDEL 318

RESULT 4
Q26268
ID Q26268 PRELIMINARY; PRT; 405 AA.
AC Q26268;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)

```



```
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidae;
OC Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93098937; PubMed=1463604;
RA Kennedy T.E., Kuhl D., Barzilai A., Sweatt J.D., Kandel E.R.;
RT "Long-term sensitization training in Aplysia leads to an increase in
RT calreticulin, a major presynaptic calcium-binding protein.";
RL Neuron 9:1013-1024(1992).
DR EMBL: S51239; AAB24569.1; -
DR InterPro: IPR000886; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PRODOM: PD001866; -; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 405 AA; 46738 MW; 14CAA201840D1D69 CRC64;

Query Match 75.0%; Score 1164; DB 5; Length 405;
Best Local Similarity 72.9%; Pred. No. 3.7e-76;
Matches 204; Conservative 29; Mismatches 39; Indels 8; Gaps 2;

QY 1 GPGTKKVVHVIYFKGKLVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
DB 134 GPGTKKVVHVIYFKGKLVINKDIRCKDDVFSHTLYTLIVRPDNTYEVKIDNEKESGDL 193

QY 61 DDWDFLPKKIKDPDASKPDWDERAKIDDDPTDSKPEDWDKPEHIPDPDAKKPDWDEEM 120
DB 194 ADWDFLPKTIPTDPAKKPDWDERAKIDDDPTDTPKPEDWDKPEHIPDPDAKKPDWDEEM 253

QY 121 DGEWEPVIONPEYKGEWKPRQIDNPYKGTWIIHPEIDNPYSPDPSIYAYDNFVGLGLD 180
DB 254 DGEWEPVIONPEYKGEWKPRQIDNPYKGTWIIHPEIDNPYSPDPSIYAYDNFVGLGLD 313

QY 181 LMQVKSCTIFDNFLITNDEAYAEFGNETGWYTKAAEKQMKDKODEORLKEEEDKKRK 240
DB 314 LMQVKSCTIFDNFLITNDEAYAEFGNETGWYTKAAEKQMKDKODEORLKEEEDKKRK 373

QY 241 EEEAEDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
DB 374 EEEAEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 405

RESULT 5
Q9PUC1 ID Q9PUC1 PRELIMINARY; PRT; 417 AA.
AC Q9PUC1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALRETICULIN.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubinstein A.L., Lee D., Luo R., Henion P.D., Halpern M.E.;
RT "Genes Dependent on Zebrafish Cyclops Function Identified by AFLP
RT Differential Gene Expression Screen.";
RL Genesis 0:0-0(1999).
DR EMBL: AF195882; AAF13700.1; -
DR InterPro: IPR000886; -
DR InterPro: IPR001580; -
```

```
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 417 AA; 48723 MW; 2000C5B4004699B6 CRC64;

Query Match 73.1%; Score 1136; DB 13; Length 417;
Best Local Similarity 70.5%; Pred. No. 4e-74;
Matches 198; Conservative 42; Mismatches 39; Indels 2; Gaps 2;

QY 1 GPGTKKVVHVIYFKGKLVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
DB 138 GYSTKVKVHVIYFKGKLVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNEKVESGSL 197

QY 61 DDWDFLPKKIKDPDASKPDWDERAKIDDDPTDSKPEDWDKPEHIPDPDAKKPDWDEEM 120
DB 198 EDWDFLPKKIKDPDASKPDWDERAKIDDDPTDTPKPEDWDKPEHIPDPDAKKPDWDEEM 257

QY 121 DGEWEPVIONPEYKGEWKPRQIDNPYKGTWIIHPEIDNPYSPDPSIYAYDNFVGLGLD 180
DB 258 DGEWEPVIONPEYKGEWKPRQIDNPYKGTWIIHPEIDNPYSPDPSIYAYDNFVGLGLD 317

QY 181 LMQVKSCTIFDNFLITNDEAYAEFGNETGWYTKAAEKQMKDKODEORLKEEEDKKRK 240
DB 318 LMQVKSCTIFDNFLITNDEAYAEFGNETGWYTKAAEKQMKDKODEORLKEEEDKKRK 376

QY 241 EEEAEDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
DB 377 DNEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 417

RESULT 6
Q91711 ID Q91711 PRELIMINARY; PRT; 343 AA.
AC Q91711
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CALRETICULIN (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=93074997; PubMed=1445218;
RA Treves S., Zorzato F., Pozzan T.;
RT "Identification of calreticulin isoforms in the central nervous
RT system.";
RL Biochem. J. 287:579-581(1992).
DR EMBL: X67598; CAA47867.1; -
DR InterPro: IPR001580; -
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 3.
DR NON_TER 1
FT NON_TER 343
RP SEQUENCE 343 AA; 40105 MW; 3E7DDAFA33B91DE1 CRC64;

Query Match 72.1%; Score 1120; DB 13; Length 343;
Best Local Similarity 83.8%; Pred. No. 4.5e-73;
Matches 192; Conservative 21; Mismatches 16; Indels 0; Gaps 0;

QY 1 GPGTKKVVHVIYFKGKLVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
DB 138 GYSTKVKVHVIYFKGKLVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNEKVESGSL 197
```

Db	110	GPTTKVHVHVFQYKKKMLQINKDI	RCCKDOSFTHLTYLTVRPDNTYEVK	IDNSKVESGSLE	169
Qy	61	DDWDFLPPKKIKDPDASKDPEDWDERAKIDDP	TSKDPEDWDKPEHIPPDPDAKPPEDWDEEM	120	
Db	170	EDMDFLPPKKIKDPEAKKPDWDERPKIDDP	EDKPKEDWKEKPEHIPPDPDAKPPEDWDEEM	229	
Qy	121	DGEWEPVIONPEYKGEWKPRQIDNPDPYKCTW	IHPDINPEYSPDPSIYAYDNFVGLGLD	180	
Db	230	DGEWEPVIONPDLQGEWKPRQIDNPDPYKGTW	IHPDINPEYPTDPTLSYESFGVIGLD	289	
Qy	181	LWQVKSITFDNLIITNDEAYAEFEGNETWGT	VKAAEKOMKQDEQR	229	
Db	290	LWQVKSITFDNLFMTNDEKXAEYNETWGT	VTKAEKKMKQDEDR	338	
RESULT	7				
Q9U5G0		PRELIMINARY;	PRT;	321	AA.
AC	Q9U5G0;				
ID	Q9U5G0;				
DT	01-MAY-2000 (TrEMBLrel. 13, Created)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update).				
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)				
DE	CALRETICULIN (FRAGMENT).				
OS	Eptareticus burgeri (Inshore hagfish).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;				
OC	Myxiniidae; Eptaretinae; Eptaretus.				
OX	NCBI_TaxID=7764;				
RP	[1]				
RN	SEQUENCE FROM N.A.				
RC	TISSUE=LIVER;				
RC	MDLINE=20063780; PubMed=10594174;				
RL	Kuraku S., Hoshiyama D., Katoh K., Suga H., Miyata T.;				
RT	"Monophyly of lampreys and hagfishes supported by nuclear DNA-coded				
RT	genes,"				
RL	J. Mol. Evol. 49:729-735(1999).				
DR	ENBL; AB025323; BAA88476.1; -.				
DR	InterPro: IPR000886; -.				
DR	InterPro: IPR001580; -.				
DR	Pfam: PF00262; calreticulin.1.				
DR	PRINTS: PR00626; CALRETICULIN.				
DR	ProDom: PD001866; -; 1.				
DR	PROSITE; PS00804; CALRETICULIN_2; 1.				
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.				
DR	PROSITE; PS00014; ER_TARGET; UNKNOWN_1.				
FT	NON_TER	1			
FT	SEQUENCE	321	AA; 37367	MR; 6E8DFA98D42F7AEF	CRC64;
Query Match		70.5%;	Score 1095.5;	DB 13;	Length 321;
Best Local Similarity		67.6%;	Pred. No. 2.4e-71;		
Matches 194;	Conservative	38;	Mismatches	46;	Indels 9; Gaps
Qy	1	GPCTKKVHVTFYKGNVLINKDIRCKDDEFTHLTYLTVRPDNTYEVKIDNSQVESGSLE	60		
Db	37	GYSTKKVHVTLNLSKGNHLIKKCKDEMTHLTYLMLYDPQTVYEVVDQSKIESGSLT	96		
Qy	61	DDWDFLPPKKIKDPDASKDPEDWDERAKIDDP	TSKDPEDWDKPEHIPPDPDAKPPEDWDEEM	120	
Db	97	EDMDFLPPKKIKDPEAKKPDWDERPKIDDP	EDKPKEDWKEKPEHIPPDPDAKPPEDWDEEM	156	
Qy	121	DGEWEPVIONPEYKGEWKPRQIDNPDPYKGTW	IHPDINPEYSPDPSIYAYDNFVGLGLD	180	
Db	157	DGEWEPVIONPDLQGEWKPRQIDNPDPYKGTW	IHPDINPEYPTDPTLSYESFGVIGLD	216	
Qy	181	LWQVKSITFDNLIITNDEAYAEFEGNETWGT	VKAAEKOMKQDEQR	229	
Db	217	LWQVKSITFDNLIITNDEKXAEYNETWGT	VTKAEKKMKQDEDR	289	
Qy	240	-----KEEEAEADKDEDEDEDEDEDEKE-EDEEDYVPGAKDEL	280		
Db	277	DGKGDEEDADEDDDDDDDEEEDKAGESPVEETP--TKDEL	321		

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 GN CALRETICULIN PRECURSOR.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Susan J.M., Just M.L., Lennarz W.J.;
 RT "Cloning and Characterization of AlphaP Integrin and Calreticulin in
 RT Embryos of the Sea Urchin."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177915; AAD55725.1; -;
 DR InterPro; IPR000886; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; calreticulin;
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 421 CALRETICULIN.
 SQ SEQUENCE 421 AA; 48822 MW; 172C664F59F41F93 CRC64;

 Query Match 67.6%; Score 1049.5; DB 5; Length 421;
 Best Local Similarity 64.2%; Pred. No. 6.5e-68;
 Matches 185; Conservative 47; Mismatches 45; Indels 11; Gaps 3;

 QY 1 GPGTKVHVIFNYKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
 Db 137 GPGTKVHVIFNYKGNVLINKIRCKDDETHLYTLIVKSDNSYEVRIIDNEKAQAGNLE 196

 QY 61 DDWDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIIPDDAKKPEDWDEEM 120
 Db 197 EDWDFLPKMIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIIPDDAKKPEDWDEEM 256

 QY 121 DGWEPPVIONPEYKGMKPRQIDNPYKGTWHPIDNPYSPDPSIAYDNFVGLGLD 180
 Db 257 DGWEPPVIONPEYKGMKPRQIDNPYKGTWHPIDNPYSPDPSIAYDNFVGLGLD 316

 QY 181 LMQVKSITFDNFLTINDEAYAEFGNETWGTWTAABKQKQDEBQRLKEEEDKKRK 240
 Db 317 LMQVKSITFDNFLTINDEAYAEFGNETWGTWTAABKQKQDEBQRLKEEEDKKRK 376

 QY 241 EE--EAEKDEDDKDEDEED-----EEDKEDEEDVPGQAKDEL 280
 Db 377 EEGGDDEEGDEDDKDEDEED-----EEDKEDEEDVPGQAKDEL 421

 RESULT 10
 O76961 PRELIMINARY; PRT; 403 AA.
 AC O76961;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 GN CALRETICULIN PRECURSOR.
 OS Necator americanus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Ancylostomatidae; Ancylostomatidae; Bunostomatinae; Necator.
 OX NCBI_TaxID=51031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pritchard D.I., Brown A., McElroy P., Loukas A., Girdwood K.,
 RA Berry C., Fullkrug R., Beck E.;
 RT "Calreticulin is a hookworm allergen."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ006790; CAA07254.1; -;
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; calreticulin;
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 KW Signal; Allergen.
 FT SIGNAL 1 16 POTENTIAL.
 SQ SEQUENCE 403 AA; 46833 MW; 21F38B0515487B6F CRC64;

 Query Match 66.5%; Score 1033; DB 5; Length 403;
 Best Local Similarity 68.5%; Pred. No. 9.4e-67;
 Matches 185; Conservative 30; Mismatches 53; Indels 2; Gaps 2;

 QY 1 GPGTKVHVIFNYKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60
 Db 135 GP-TKKVHDIFSYKGNHLIKKIRCKDDELTHLYTLILNPDNTYEVQIDGKVESGELE 193

 QY 61 DDWDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIIPDDAKKPEDWDEEM 120
 Db 194 SDWDLFPKKIKDPDASKPEDWDEREYIDDDAKKPEDWPKPEHIIPDDAKKPEDWDEEM 253

 QY 121 DGWEPPVIONPEYKGMKPRQIDNPYKGTWHPIDNPYSPDPSIAYDNFVGLGLD 180
 Db 254 DGWEPPVIONPEYKGMKPRQIDNPYKGTWHPIDNPYSPDPSIAYDNFVGLGLD 313

 QY 181 LMQVKSITFDNFLTINDEAYAEFGNETWGTWTAABKQKQDEBQRLKEEEDKKRK 240
 Db 314 LMQVKSITFDNFLTINDEAYAEFGNETWGTWTAABKQKQDEBQRLKEEEDKKRK 372

 QY 241 EEEAEKDEDDKDEDEEDKEDEEEDKEE 270
 Db 373 EEEKKKKEEKEEKEEKEEKEEKEE 402

 RESULT 11
 Q9U916 PRELIMINARY; PRT; 406 AA.
 AC Q9U916;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 GN CALRETICULIN
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R;
 RA Dodo K., Sakoyama Y., Gamo S.;
 RT "Drosophila melanogaster calreticulin for mRNA."
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB000718; BAA85379.1; -;
 DR FlyBase; FBgn0005585; Crc.
 DR InterPro; IPR000886; -;
 DR InterPro; IPR001580; -;
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN
 DR PROSITE; PS00803; CALRETICULIN_1; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 SQ SEQUENCE 406 AA; 46809 MW; 68BA49A6B81CC427 CRC64;

۵۰

OC Magnoliophyta; Lilliolepis; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96096524; PubMed=8522179;
RA Kwiatkowski L.A., Zielinska-Kwiatkowska A.G., Migdalski A.,
RA Kleczkowski L.A., Wasilewska L.D.;
RC STRAIN=VAR MERIT; TISSUE=ROOT TIP;
RA Napier R.M., Trueman S., Henderson J., Boyce J.M., Hawes C.R.,
RA Fricker M.D., Venis M.A.;
RL J. Exp. Bot. 46:1603-1613(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96309381; PubMed=8704156;
RA Desselhaus T., Hagel C., Loefer H., Kranz E.;
RC "Isolation of a full-length cDNA encoding calreticulin from a PCR
RT library of in vitro zygotes of maize.";
RL Plant Mol. Biol. 31:23-34(1996).
DR EMBL: Z46772; CAA86728.1; -;
DR EMBL: X89813; CAA61939.1; -;
DR Mendel: 11228; Zeama:1166;11228.
DR InterPro: IPR000886; -;
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PRODOM: PD001866; -; 1.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW Signal; Calcium-binding.
FT SIGNAL 1 25
FT CHAIN 26 421
FT CHAIN 26 421 CALRETICULIN.
SQ SEQUENCE 421 AA; 48012 MW; 5AAE02B77ED3126D CRC64;

Query Match 54.3%; Score 844; DB 10; Length 421;
Best Local Similarity 53.7%; Pred. No. 3.6e-53;
Matches 154; Conservative 46; Mismatches 71; Indels 16; Gaps 5;

QY 1 GPGTKKVVHVFYFNKGNVILNKIRCKDDEFTHTLTVIRPDNTYEVKIDNSQVSGSLE 60
Db 144 GYSTKKVHTLTLDGKNHLIKKDVPCETDQTLTHVYTLIRPDATYSILIDNEEKQTGSY 203

QY 61 DDMDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWK-PHIPPDPDAKPEDWDEE 119
Db 204 EHWDLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWK-PHIPPDPDAKPEDWDEE 263

QY 120 MDGEWPEPPVIONPEYKGEWKPQIDNPDKYGTWTHPEIDNPESPDPSIYAYDNFVGL 179
Db 264 EDGEWTAFTIPNPEYKGPWKQKIKNPNYQGWKAPMIDNPDKDDPYIYAFDSLKYIGI 323

QY 180 DLQVKSGLTFDNLITNDAYAEFGNETWGTWTKAAEKQMKDKQDEEQRKEEEDKKR 239
Db 324 ELQVKSGLTFDNLITNDAYAEFGNETWGTWTKAAEKQMKDKQDEEQRKEEEDAAK 379

QY 240 ---KEEEAEDEKDE 280
Db 380 GGDEDDDE 421

RESULT 15
Q41799 PRELIMINARY; PRT; 321 AA.
AC Q41799;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CALRETICULIN (FRAGMENT).
GN CRH.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Lilliolepis; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96096524; PubMed=8522179;
RA Kwiatkowski L.A., Zielinska-Kwiatkowska A.G., Migdalski A.,
RA Kleczkowski L.A., Wasilewska L.D.;
RC "Cloning of two cDNAs encoding calnexin-like and calreticulin-like
RT proteins from maize (Zea mays) leaves: identification of potential
RT calcium-binding domains.";
RL Gene 165:219-222(1995).
DR EMBL: X78057; CAA54975.1; -;
DR Mendel: 11227; Zeama:1166;11227.
DR InterPro: IPR000886; -;
DR Pfam: PF00262; calreticulin; 1.
DR PRINTS: PR00626; CALRETICULIN.
DR PROSITE: PS00803; CALRETICULIN_1; 1.
DR PROSITE: PS00804; CALRETICULIN_2; 1.
DR PROSITE: PS00805; CALRETICULIN_REPEAT; 2.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 321 AA; 36583 MW; 687CC4460ABE6AF1 CRC64;

Query Match 54.3%; Score 843.5; DB 10; Length 321;
Best Local Similarity 52.7%; Pred. No. 2.9e-53;
Matches 148; Conservative 50; Mismatches 78; Indels 5; Gaps 2;

QY 1 GPGTKKVVHVFYFNKGNVILNKIRCKDDEFTHTLTVIRPDNTYEVKIDNSQVSGSLE 60
Db 45 GYSTKKVHTLTLDGKNHLIKKDVPCETDQTLTHVYTLIRPDATYSILIDNEEKQTGSY 104

QY 61 DDMDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWK-PHIPPDPDAKPEDWDEE 119
Db 105 EHWDLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWK-PHIPPDPDAKPEDWDEE 164

QY 120 MDGEWPEPPVIONPEYKGEWKPQIDNPDKYGTWTHPEIDNPESPDPSIYAYDNFVGL 179
Db 165 EDGEWTAFTIPNPEYKGPWKQKIKNPNYQGWKAPMIDNPDKDDPYIYAFDSLKYIGI 224

QY 180 DLQVKSGLTFDNLITNDAYAEFGNETWGTWTKAAEKQMKDKQDEEQRKEEEDKKR 239
Db 225 ELQVKSGLTFDNLITNDAYAEFGNETWGTWTKAAEKQMKDKQDEEQRKEEEDAAK 280

QY 240 KEEAEDEKDE 280
Db 281 GGDEDDDE 321

Search completed: October 21, 2001, 03:10:09
Job time: 344 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:07:44 ; Search time 52.92 Seconds
(without alignments)
259.097 Million cell updates/sec

Title: US-09-807-148-4
Perfect score: 971
Sequence: 1 EPAAVFEKQFLDGDGWTSRW.....PNTVEYKIDNSQVSGSL80

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	971	100.0	417	1 A37047	calreticulin precu
2	933	96.1	416	2 JH0819	calreticulin precu
3	931	95.9	418	1 A34154	calreticulin precu
4	930	95.8	400	2 S43376	calreticulin, brai
5	930	95.8	416	1 S06763	calreticulin precu
6	835	86.0	411	2 S29129	calreticulin precu
7	804	82.8	405	1 JH0795	calreticulin precu
8	797	82.1	384	2 S29130	calreticulin (clon
9	796	82.0	419	2 S71343	calreticulin precu
10	699	72.0	406	2 A56637	calreticulin homol
11	619.5	63.8	395	2 S25851	calreticulin precu
12	542.5	55.9	336	2 A32507	41K larval antigen
13	539	55.5	412	2 T05703	calreticulin - bar
14	539	55.5	415	2 T05705	calreticulin - bar
15	536	55.2	421	2 S58170	calreticulin precu
16	531	54.7	416	2 T14554	calreticulin - bee
17	527	54.3	415	2 T10172	calreticulin - cas
18	522.5	53.8	421	2 S36799	calreticulin precu
19	522	53.8	416	2 T16968	calreticulin call
20	517	53.2	389	2 T03691	calreticulin - com
21	512	52.7	425	2 C96605	calreticulin (Crt1
22	510	52.5	444	2 H86224	hypothetical prote
23	496	51.1	393	1 A48573	calreticulin autoa
24	347	35.7	422	2 T07841	probable calreticu
25	293	30.2	591	2 B54354	calnexin precursor
26	290	29.9	591	2 C54354	calnexin precursor
27	290	29.9	592	2 I53260	calnexin - human
28	290	29.9	593	1 A37273	calnexin precursor
29	288	29.7	592	2 A46673	calnexin precursor

30	270	27.8	611	2 A53418	calmeglin precursor
31	262	27.0	622	2 S71342	calnexin precursor
32	253.5	26.1	560	2 S56142	calcium-binding pr
33	252	26.0	611	2 A54086	calnexin-t - mouse
34	248	25.5	582	2 A46637	calnexin homolog S
35	244	25.1	619	2 S40938	hypothetical prote
36	238.5	24.6	530	2 JN0597	calnexin-like prot
37	232	23.9	546	2 T06415	calnexin - soybean
38	215.5	22.2	540	2 T10892	probable calnexin
39	200.5	20.6	532	2 T49873	calnexin homolog -
40	157	16.2	29	2 E33208	calreticulin, uter
41	145	14.9	29	2 C33208	calreticulin, slow
42	144	14.8	29	2 D33208	calreticulin, brai
43	133.5	13.7	502	2 S29347	calnexin homolog Y
44	129	13.3	297	2 S70552	calnexin homolog C
45	124.5	12.8	428	2 T03251	calnexin - maize (

ALIGNMENTS

RESULT 1
A37047
calreticulin precursor - human
N;Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding pr
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence,revision 10-Sep-1999 #text_change 18-Feb-2000
C;Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075
R;McCaulliffe, D.P.; Yang, Y.S.; Wilson, J.; Sontheimer, R.D.; Capra, J.D.
J. Biol. Chem. 267, 2557-2562, 1992
A;Title: The 5'-flanking region of the human calreticulin gene shares homology with t
A;Reference number: A42330; MUID:92129342
A;Accession: A42330
A;Molecule type: DNA
A;Residues: 1-417 <MC2>
A;Note: sequence extracted from NCBI backbone (NCBIN:78537, NCBIP:78536)
R;McCaulliffe, D.P.; Lux, F.A.; Lieu, T.S.; Sanz, I.; Hanke, J.; Newkirk, M.M.; Bachlin
J. Clin. Invest. 85, 1379-1391, 1990
A;Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/
A;Reference number: A37047; MUID:90237213
A;Accession: A37047
A;Molecule type: mRNA
A;Residues: 1-417 <MCC>
A;Cross-references: GB:M32294; NID:g337486; PIDN:AAA36582.1; PID:g337487
R;Rokeach, L.A.; Haselby, J.A.; Meilof, J.F.; Smeenk, R.J.; Unnasch, T.R.; Greene, B.
J. Immunol. 147, 3031-3039, 1991
A;Title: Characterization of the autoantigen calreticulin.
A;Reference number: A46452; MUID:92013129
A;Accession: A46452
A;Molecule type: mRNA
A;Residues: 1-417 <ROK>
A;Cross-references: GB:M84739; NID:g179881; PIDN:AAA51916.1; PID:g179882
A;Note: sequence extracted from NCBI backbone (NCBIN:60749, NCBIP:60750)
R;Lieu, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.
J. Clin. Invest. 82, 96-101, 1988
A;Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence
A;Reference number: A28812; MUID:88273610
A;Accession: A28812
A;Molecule type: protein
A;Residues: 18-41 <LIE>
A;Note: 18-Ala was also found
R;Dupuis, M.; Schaefer, E.; Krause, K.H.; Tschopp, J.
J. Exp. Med. 177, 1-7, 1993
A;Title: The calcium-binding protein calreticulin is a major constituent of lytic gra
A;Reference number: PH1525; MUID:93115648
A;Accession: PH1525
A;Molecule type: protein
A;Residues: 18-27 <DUP>
A;Experimental source: LAK cell
R;Rojiani, M.V.; Finlay, B.B.; Gray, V.; Dedhar, S.
Biochemistry 30, 9859-9866, 1991
A;Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (c

A:Reference number: A40346; MUID:92002034

A:Accession: A40346

A:Molecule type: protein

A:Residues: 18-34, 'R' <ROQ>

R:Krause, K.H.; Simmerman, H.K.B.; Jones, L.R.; Campbell, K.P.

Biochem. J. 270, 545-548, 1990

A:Title: Sequence similarity of calreticulin with a Ca(2+)-binding protein that co-purifies with the Golgi apparatus

A:Reference number: S11475; MUID:90380058

A:Accession: S11475

A:Molecule type: protein

A:Residues: 18-32 <KRA>

R:Lamerdin, J.; McCreedy, P.; Stilwagen, S.; Ramirez, M.; Carrano, A.

submitted to the EMBL Data Library, November 1996

A:Description: Characterization by genomic sequence analysis of a gene-rich 111 kb region

A:Reference number: 222906

A:Accession: T45075

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-417 <LAMS>

A:Cross-references: EMBL:AD000092; PIDN:AA51176.1

A:Experimental source: cell line 5HL2-B; fibroblast

C:Comment: Autoantibodies specific for this protein are found in Sjogren's syndrome and in other autoimmune diseases

C:Genetics:

A:Gene: GDB:CALR

A:Cross-references: GDB:125179; OMIM:109091

A:Map position: 19p13.3-19p13.2

A:Introns: 31/1; 65/1; 133/1; 164/3; 272/3; 320/3; 351/3

A:Note: CRTC

C:Superfamily: calreticulin

C:Keywords: calcium binding; integrin binding

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-417/Product: calreticulin #status predicted <MAT>

F:414-417/Region: endoplasmic reticulum retention signal

Query Match 100.0%; Score 971; DB 1; Length 417;

Best Local Similarity 100.0%; Pred. No. 3.3e-80;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVKEQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSODARFYAL 60

DB 18 EPAYVKEQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSODARFYAL 77

QY 61 SASPEFNSKNGQTLVVOFTVKHEQNDICDGGYVKLFPNSLDQTDHMGDSYINMEGPDIC 120

DB 78 SASPEFNSKNGQTLVVOFTVKHEQNDICDGGYVKLFPNSLDQTDHMGDSYINMEGPDIC 137

QY 121 GPGTKVHVIFNFKGNVLINKDIRCKDEFFHLYTLIVRPDNTYEVKIDNSQVSGSLE 180

DB 138 GPGTKVHVIFNFKGNVLINKDIRCKDEFFHLYTLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 2

JH0819

calreticulin precursor - rat

N:Alternate names: calcium-binding protein 3

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000

C:Accession: JH0819; A49176; S11205; PC1109; S45036; S04867; S39372; A34473; S13045

R:Nakamura, M.; Moriya, M.; Baba, T.; Michikawa, Y.; Yamanobe, T.; Arai, K.; Okinaga, S.

Exp. Cell Res. 205, 101-110, 1993

A:Title: An endoplasmic reticulum protein, calreticulin, is transported into the acrosome

A:Reference number: A49176; MUID:93202172

A:Accession: JH0819

A:Molecule type: mRNA

A:Residues: 1-416 <NAK>

A:Cross-references: GB:D78308; NID:g1089798; PIDN:BAAL1345.1; PID:g1845572

A:Accession: A49176

A:Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-416 <NAK>

A:Cross-references: GB:D78308; NID:g1089798; PIDN:BAAL1345.1; PID:g1845572

A:Experimental source: Sprague-Dawley, spermatogenic cells

A:Note: sequence extracted from NCBI backbone (NCBIN:127639, NCBIP:127643)

R:Murthy, K.K.; Banville, D.; Srikanth, C.B.; Carrier, F.; Holmes, C.; Bell, A.; Patel

Nucleic Acids Res. 18, 4933, 1990

A:Title: Structural homology between the rat calreticulin gene product and the Onchoc

A:Reference number: S11205; MUID:90370496

A:Accession: S11205

A:Molecule type: mRNA

A:Residues: 1-416 <MRU>

A:Cross-references: EMBL:X53363; NID:g55854; PIDN:CAA37446.1; PID:g55855

R:Nakamura, M.; Michikawa, Y.; Baba, T.; Okinaga, S.; Arai, K.

Biochem. Biophys. Res. Commun. 186, 668-673, 1992

A:Title: Calreticulin is present in the acrosome of spermatids of rat testis.

A:Reference number: PC1109; MUID:92360010

A:Accession: PC1109

A:Molecule type: protein

A:Residues: 18-32 <NAK2>

A:Experimental source: testis, strain Sprague-Dawley

R:Soennichsen, B.; Fuellekrug, J.; van Nguyen, P.; Diekmann, W.; Robinson, D.G.; Mies

submitted to the EMBL Data Library, May 1994

A:Description: Retention and retrieval: both mechanisms cooperate to maintain calreti

A:Reference number: S45036

A:Accession: S45036

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <SOE>

A:Cross-references: EMBL:X79327; NID:g488840; PIDN:CAA55890.1; PID:g488841

R:Lone, Y.C.; Bailly, A.; Latruffe, N.

submitted to the EMBL Data Library, December 1988

A:Reference number: S04867

A:Accession: S04867

A:Molecule type: mRNA

A:Residues: 'R', 270-358, 'AAG' <LON>

A:Cross-references: EMBL:X13702; NID:g56055; PIDN:CAA31987.1; PID:g930260

A:Note: the authors designated the protein as D-beta-hydroxybutyrate dehydrogenase

R:Yokoi, T.; Nagayama, S.; Kajiwara, R.; Kawaguchi, Y.; Horiuchi, R.; Kametaki, T.

Biochim. Biophys. Acta 1158, 339-344, 1993

A:Title: Identification of protein disulfide isomerase and calreticulin as autoimmune

A:Reference number: S39371; MUID:94072621

A:Accession: S39372

A:Molecule type: protein

A:Residues: 18-23, X', 25-32 <YOK>

R:Van, P.N.; Peter, F.; Soeling, H.D.

J. Biol. Chem. 264, 17494-17501, 1989

A:Title: Four intracisternal calcium-binding glycoproteins from rat liver microsomes

active calcium sequestering rat liver vesicles.

A:Reference number: A34473; MUID:90008920

A:Accession: A34473

A:Status: preliminary

A:Molecule type: protein

A:Residues: 18-36 <VAN>

R:Treves, S.; de Mattel, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; M

Biochem. J. 271, 473-480, 1990

A:Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-stor

A:Reference number: S13045; MUID:91054414

A:Accession: S13045

A:Molecule type: protein

A:Residues: 18-29 <TRE>

C:Superfamily: calreticulin

C:Keywords: calcium binding; glycoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-416/Product: calreticulin #status experimental <MAT>

F:204-212/Region: nuclear location signal

F:413-416/Region: endoplasmic reticulum retention signal

F:344/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 96.1%; Score 933; DB 2; Length 416;

Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPAYVKEQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSODARFYAL 60

DB 18 DPALVKEQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSODARFYAL 77

Qy 61 SASPEPSNKGQTLLVQFTVKHKEQNIDCGGYVKLFNPNSLDQDTDMHGDSYNIIMFGPDIC 120
 ||| |
Dd 78 SARPEPSSNKGTLLVQFTVKHKEQNIDCGGYVKLFCGGLDQKMDMGDSEYNIMFGPDIC 137

Qy 121 GPCTKKVHVIFNYKGNVLINKDIRCKDEFFTHLYTLIVRPDNTYEVDINSQVESGSLE 180
 ||| |
Dd 138 GPGTKKVHVIFNYKGNVLINKDIRCKDEFFTHLYTLIVRPDNTYEVDINSQVESGSLE 197

RESULT 3

A34154
calreticulin precursor, skeletal muscle - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34154; S13047
R:Filiegel, L.; Burns, K.; MacLennan, D.H.; Reithmeier, R.A.F.; Michalak, M.
J. Biol. Chem. 264, 21522-21528, 1989
A>Title: Molecular cloning of the high affinity calcium-binding protein (calreticulin) cDNA from rabbit skeletal muscle.
A:Reference number: A34154; MUID:90094320
A:Accession: A34154
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <FRT>
A:Cross-references: GB:J05138; NID:g164858; PIDN:AAA31188.1; PID:g164859
R:Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meldolesi, J. Biochemistry International 19, 1-17, 1988
Biochem. J. 271, 473-480, 1990
A>Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage organelles.
A:Reference number: S13045; MUID:91054414
A:Accession: S13047
A:Molecule type: protein
A:Residues: 19-32 <TRE>
C:Superfamily: calreticulin
C:Keywords: skeletal muscle
F:1-17/Domains: signal sequence #status predicted <SIG>
F:415-418/Region: endoplasmic reticulum retention signal

```

Query Match          95.9%; Score 931.; DB 1; Length 418;
Best Local Similarity 95.6%; Pred. No. 1.3e-76;
Matches 172; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 EPAYFYKEQFLDGGWTSRWIESKHKHSDFKFLVSSGKFGYGDDEKKGLOTSDQAREYAL 60
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18 EPVYFYKEQFLDGGWTSRWIESKHKHSDFKFLVSSGKFGYGDDEKKGLOTSDQAREYAL 77
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 SASPEPSNKGQTLVQVFTVKVHEONIDCGGYVKLRFNSLDQTDMDHGDSEYNTMFGPDIC 120
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78 SARPEPSNKGQTLVQVFTVKVHEONIDCGGYVKLRFNSLDQTDMDHGDSEYNTMFGPDIC 137
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GPCTKKVHVIFYNKGKNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 GPCTKKVHVIFYNKGKNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 197

RESULT 4
S43376
calreticulin, brain isoform 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Oct-1994 #sequence_revision 23-Mar-1995 #text_change 07-May-1999
C:Accession: S43376; S36801
R:Batsuoka, K.; Seta, K.; Yamakawa, Y.; Okuyama, T.; Shinoda, T.; Isobe, T.
Biochem. J. 298, 435-442, 1994
A:Title: Covalent structure of bovine brain calreticulin.
A:Reference number: S43376; MUID:94183174
A:Accession: S43376
A:Molecule type: protein
A:Residues: 1-400 <MAT>
A:Experimental source: brain
R:Liu, N.; Fine, R.E.; Johnson, R.J.
Biochim. Biophys. Acta 1202, 70-76, 1993
A:Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin.
A:Reference number: S36799; MUID:93385184

```

A:Accession: S36801
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 45-63, 'E', 65-83 <LTIU>
A:Experimental source: brain, clone 8.1
C:Superfamily: calreticulin
C:Keywords: calcium binding; glycoprotein
F:397-400/Region: endoplasmic reticulum retention signal
F:120-146/Disulfide bonds: #status experimental
F:162/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 95.88; Score 930; DB 2; Length 400;
Best Local Similarity 95.08; Pred. No. 1.6e-76;
Matches 171; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 EPAYFYEQFDLGGWTSRWIESKHKSDPKFVLSSGKFYGDDEKDKGLQTSQDARFYAL 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 DPTVFYEQFDLGGWTSRWIESKHKSDPKFVLSSGKFYGDDEKDKGLQTSQDARFYAL 60

Qy 61 SASPEFSSNKGQTLVVOFTVKVHEQNI DCGGYVKLFPSLDQTDHMGDSEYNIMFGPDIC 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 SARPEFSSNKGQTLVVOFTVKVHEQNI DCGGYVKLFPSLDQTDHMGDSEYNIMFGPDIC 120

Qy 121 GPGTKKVVHVFNYKGNVLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKINDSQVESGSLE 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 GPGTKKVVHVFNYKGNVLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKINDSQVESGSLE 180

RESULT 5
S06763
calreticulin precursor - mouse
N:Alternate names: 55K calcium-binding reticuloplasmin; calregulin
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S06763; JCI444; PCL1233; A57498
R:Smith, M.J.; Koch, G.L.E.
EMBO J. 8, 3581-3586, 1989
A:Title: Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a
A:Reference number: S06763; MUID:50059955
A:Accession: S06763
A:Molecule type: DNA
A:Residues: 1-416 <SMI>
A:Cross-references: EMBL:X14926; NID:g50567; PIDN:CAA33053.1; PID:g50568
R:Mazzarella, R.A.; Gold, P.; Cunningham, M.; Green, M.
Gene 120, 217-225, 1992
A:Title: Determination of the sequence of an expressible cDNA clone encoding Erp60/ca
A:Reference number: JCI444; MUID:93013037
A:Accession: JCI444
A:Molecule type: mRNA
A:Residues: 1-416 <MAZ>
A:Cross-references: GB:M92988; NID:g193084; PIDN:AAA37569.1; PID:g193085
A:Accession: PCL1233
A:Molecule type: protein
A:Residues: 18-41 <MAZ>
R:White, T.K.; Zhu, Q.; Tanzer, M.L.
J. Biol. Chem. 270, 15926-15929, 1995
A:Title: Cell surface calreticulin is a putative mannoside lectin which triggers mous
A:Reference number: A57498; MUID:95332280
A:Accession: A57498
A:Status: preliminary
A:Molecule type: protein
A:Residues: 74-80; 142-151; 186-193 <WHI>
C:Superfamily: calreticulin
C:Keywords: calcium binding
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-416/Product: calregulin #status experimental <MAT>
F:413-416/Region: endoplasmic reticulum retention signal

Query Match 95.88; Score 930; DB 1; Length 416;
Best Local Similarity 94.4%; Pred. No. 1.6e-76;
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 EPAVYKEQFLDGDGWTNRWIESKHKSDGKFKVLSGKFGYDEEDKGLQTSQDARFYAL 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 18 DPAVYKEQFLDGDGWTNRWIESKHKSDGKFKVLSGKFGYDEEDKGLQTSQDARFYAL 77
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 SASFEFSSKNGQTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDIC 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 78 SAKFEFSSKNGQTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDIC 137
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPNTYEVKIDNSQVSGSLE 180
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 138 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPNTYEVKIDNSQVSGSLE 197
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
RESULT 6
S29129
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: S29129
R:Treves, S.; Zorzato, F.; Pozzan, T.
Biochem. J. 287, 579-581, 1992
A:Title: Identification of calreticulin isoforms in the central nervous system.
A:Reference number: S29129; MUID:93074997
A:Accession: S29129
A:Molecule type: mRNA
A:Residues: 1-411 <TRE>
A:Cross-references: EMBL:X67597; NID:g64608; PIDN:CAA47866.1; PID:g64609
A:Superfamily: calreticulin
C:Keywords: glycoprotein
F:1-12/Domain: signal sequence (fragment) #status predicted <SIG>
F:13-411/Product: calreticulin #status predicted <MAT>
F:408-411/Region: endoplasmic reticulum retention signal
F:339/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 86.0%; Score 835; DB 2; Length 411;
Best Local Similarity 83.9%; Pred. No. 6.2e-68;
Matches 151; Conservative 15; Mismatches 14; Indels 0; Gaps 0;

QY 1 EPAVYKEQFLDGDGWTNRWIESKHKSDGKFKVLSGKFGYDEEDKGLQTSQDARFYAL 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 13 EPAVYKEQFLDGDGWTNRWIESKHKSDGKFKVLSGKFGYDEEDKGLQTSQDARFYAM 72
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 SASFEFSSKNGQTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDIC 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 73 SSRFDSFNKQTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDIC 132
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPNTYEVKIDNSQVSGSLE 180
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 133 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPNTYEVKIDNSQVSGSLE 192
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
RESULT 7
JH0795
calreticulin precursor - California sea hare
N:Alternate names: protein 407
C:Species: Aplysia californica (California sea hare)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0795; B31409; F60977
R:Kennedy, T.E.; Kuhl, D.; Barzilai, A.; Sweatt, J.D.; Kandel, E.R.
Neuron 9, 1013-1024, 1992
A:Title: Long-term sensitization training in aplysia leads to an increase in calreticulin
A:Reference number: JH0795; MUID:93098937
A:Accession: JH0795
A:Molecule type: mRNA
A:Residues: 1-405 <KEN>
A:Cross-references: GB:S51239; NID:g262053; PIDN:AA824569.1; PID:g262054
A:Experimental source: abdominal ganglion and antral nervous system
R:Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.; Sweatt, J.D.
Proc. Natl. Acad. Sci. U.S.A. 85, 7008-7012, 1988
A:Title: Sequencing of proteins from two-dimensional gels by using in situ digestion and
tion in Aplysia.

```

```

A:Reference number: A94207; MUID:88320566
A:Accession: B31409
A:Molecule type: protein
A:Residues: 'X', 17-28, 'X', 30-31 <KE2>
R:Sweatt, J.D.; Kennedy, T.E.; Wager-Smith, K.; Gawinowicz, M.A.; Barzilai, A.; Karl,
Electrophoresis 10, 152-157, 1989
A:Title: Development of a database of amino acid sequences for proteins identified an
A:Reference number: A60977; MUID:89276264
A:Accession: F60977
A:Molecule type: protein
A:Residues: 'X', 17-28, 'X', 30-31 <SWE>
C:Superfamily: calreticulin
C:Keywords: calcium binding; endoplasmic reticulum
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-405/Product: calreticulin #status experimental <MAT>
F:402-405/Region: endoplasmic reticulum retention signal

Query Match 82.8%; Score 804; DB 1; Length 405;
Best Local Similarity 79.4%; Pred. No. 3.9e-65;
Matches 143; Conservative 18; Mismatches 17; Indels 2; Gaps 1;

QY 1 EPAVYKEQFLDGDGWTNRWIESKHKSDGKFKVLSGKFGYDEEDKGLQTSQDARFYAL 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 16 DPTVYFKEEF--GDDWAERWVESKHKSDLGKFKVLTAGKFGYDAEKDKGIGTQTSQDARFYGL 73
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 SASFEFSSKNGQTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDIC 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 74 SAKFKFSNEGKTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDIC 133
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPNTYEVKIDNSQVSGSLE 180
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 134 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLYTLIVRPNTYEVKIDNEKAESGDLE 193
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
RESULT 8
S29130
calreticulin (clone 8) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: S29130; T01068
R:Treves, S.; Zorzato, F.; Pozzan, T.
Biochem. J. 287, 579-581, 1992
A:Title: Identification of calreticulin isoforms in the central nervous system.
A:Reference number: S29129; MUID:93074997
A:Accession: S29130
A:Molecule type: mRNA
A:Residues: 1-384 <TRE>
A:Cross-references: EMBL:X67598
A:Accession: T01068
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-339, 'XTGR' <TRW>
A:Cross-references: EMBL:X67598; NID:g64610; PIDN:CAA47867.1; PID:g64611
A:Experimental source: CNS
C:Superfamily: calreticulin
C:Keywords: glycoprotein
F:381-384/Region: endoplasmic reticulum retention signal
F:316/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 797; DB 2; Length 384;
Best Local Similarity 85.2%; Pred. No. 1.6e-64;
Matches 144; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 12 DGDGWTNRWIESKHKSDGKFKVLSGKFGYDEEDKGLQTSQDARFYALSASFEPSSNGK 71
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 DGDGWTNRWIESKHKSDGKFKVLSGKFGYDEEDKGLQTSQDARFYAMSSRFESFNKD 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 72 QTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDICGPGTKKVVHVF 131
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 QTLVQVFTVKHEQNDICGGYVKLFPNSLDQTMHGDSEYNIMFGPDICGPGTKKVVHVF 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



```
RESULT 12
A32507
41K larval antigen - nematode (Onchocerca volvulus) (fragment)
C:Species: Onchocerca volvulus
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 12-Apr-1995
C:Accession: A32507; A28813
R:Unasch, T.R.; Gallin, M.Y.; Soboslay, P.T.; Erttmann, K.D.; Greene, B.M.
J. Clin. Invest. 82, 262-269, 1988
A:Title: Isolation and characterization of expression cDNA clones encoding antigens of O.
A:Reference number: A92769; MUID:88273584
A:Accession: A32507
A:Molecule type: mRNA
A:Residues: 1-336 <UNN>
C:Superfamily: calreticulin

Query Match 55.9%; Score 542.5; DB 2; Length 336;
Best Local Similarity 68.5%; Pred. No. 1.3e-41;
Matches 98; Conservative 22; Mismatches 22; Indels 1; Gaps 1;

QY 39 FYGDEPKDGLQTSQDARYVALSASFE-PFSNKGOTLVQFTVKHEQNTDCKDDFTHLYTL 97
Db 1 FYGDVAKDKGLTKTQDAKFSISGAKDKSFSNKGSLVIFQFSVKHEQDIDCGGYVKLMA 60

QY 98 NSLDQDMHGDSEYNIMFGPDICGPGTKKVVHIFNYKGNVLINKDIRCKDDFTHLYTL 157
Db 61 SDVNLEDSHGETPYHIMFGPDICGPGTKKVVHIFNYKGNVLINKDIRCKDDVFTHLTYL 120

QY 158 IVRPDNTYEVKIDNSOVESGSL 180
Db 121 IVNSDNTYEVQIDGKAESGELE 143

RESULT 13
T05703
calreticulin - barley (fragment)
C:Species: Hordeum vulgare (barley)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T05703
R:Chen, F.; Hayes, P.M.; Mulroony, D.; Pan, A.
Plant Cell 6, 835-843, 1994
A:Title: Identification and characterization of cDNA clones encoding plant calreticulin
A:Reference number: Z15422; MUID:94339696
A:Accession: T05703
A:Molecule type: mRNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-412 <CHE>
A:Cross-references: EMBL:L27348; NID:g439585; PIDN:AAA32948.1; PID:g439586
A:Experimental source: cv. Morex, ovary
C:Genetics:
A:Gene: CRH1
A:Map position: 1M
C:Superfamily: calreticulin
C:Keywords: calcium binding

Query Match 55.5%; Score 539; DB 2; Length 412;
Best Local Similarity 56.4%; Pred. No. 3.5e-41;
Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTNRWIESKHSD---FGKFLVSSGKFGYGDDEKGLQTSQDARFYAL 60
Db 20 VFQEKFF--EDGWESRWKSEWKNMAGWNHTSGKWHGDAE-DKGILQTSSEYRFFAI 76

QY 61 SASFEPSNKGOTLVQFTVKHEQNTDCCGGYVKLFPNSLDQDMHGDSEYNIMFGPDIC 120
Db 77 SAEYPEFSNKRDTLVQFTVKHEQKLDCCGGYVKLLGGVDVQDKKFGGDPYGYIMFGPDIC 136

QY 121 GPGTKKVVHIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSOVESGSL 179
Db 137 GYSTKKVHTLTNKGKNNHLLKDKVPCETDQLSHVYTLIRPDATYSILIDNEEKQTGSI 195
```

```
RESULT 14
T05705
calreticulin - barley (fragment)
C:Species: Hordeum vulgare (barley)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T05705
R:Chen, F.; Hayes, P.M.; Mulroony, D.; Pan, A.
Plant Cell 6, 835-843, 1994
A:Title: Identification and characterization of cDNA clones encoding plant calreticulin
A:Reference number: Z15422; MUID:94339696
A:Accession: T05705
A:Molecule type: mRNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-415 <CHE>
A:Cross-references: EMBL:L27349; NID:g439587; PIDN:AAA32949.1; PID:g439588
A:Experimental source: cv. Morex, ovary
C:Genetics:
A:Gene: CRH2
A:Map position: 2M
C:Superfamily: calreticulin
C:Keywords: calcium binding

Query Match 55.5%; Score 539; DB 2; Length 415;
Best Local Similarity 56.4%; Pred. No. 3.5e-41;
Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTNRWIESKHSD---FGKFLVSSGKFGYGDDEKGLQTSQDARFYAL 60
Db 23 VFQEKFF--EDGWESRWKSEWKNMAGWNHTSGKWHGDAE-DKGILQTSSEYRFFAI 79

QY 61 SASFEPSNKGOTLVQFTVKHEQNTDCCGGYVKLFPNSLDQDMHGDSEYNIMFGPDIC 120
Db 80 SAEYPEFSNKRDTLVQFTVKHEQKLDCCGGYVKLLGGVDVQDKKFGGDPYGYIMFGPDIC 139

QY 121 GPGTKKVVHIFNYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSOVESGSL 179
Db 140 GYSTKKVHTLTNKGKNNHLLKDKVPCETDQLSHVYTLIRPDATYSILIDNEEKQTGSI 198

RESULT 15
S58170
calreticulin precursor - maize
C:Species: Zea mays (maize)
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 13-Aug-1999
C:Accession: S58170; S49818
R:Dresselhaus, T.; Hagel, C.; Loerz, H.; Kranz, E.
submitted to the EMBL Data Library, July 1995
A:Description: Isolation of a cDNA encoding Calreticulin from in vitro zygotes of mai
A:Reference number: S58170
A:Accession: S58170
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-421 <DRE>
A:Cross-references: EMBL:X89813; NID:g927571; PIDN:CAA61939.1; PID:g927572
R:Napier, R.M.; Trueman, S.; Henderson, J.; Boyce, J.M.; Hawes, C.R.; Fricker, M.D.;
submitted to the EMBL Data Library, November 1994
A:Description: Purification and sequencing of calreticulin from maize and evidence for
A:Reference number: S49818
A:Accession: S49818
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-421 <NAP>
A:Cross-references: EMBL:Z46772; NID:g577611; PIDN:CAA86728.1; PID:g577612
C:Genetics:
A:Gene: CRT1
C:Superfamily: calreticulin
C:Keywords: calcium binding
F.1-25/Domain: signal sequence #status predicted <SIC>
F.418-421/Region: endoplasmic reticulum retention signal
```

Query Match 55.2%; Score 536; DB 2; Length 421;
Best Local Similarity 56.4%; Pred. No. 6.7e-41;
Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY	4	VYFKEQFLDGDGWTSRWIESKHKSD---FGKFLVSSGKFGYGDDEKDKGLQTSQDAREYAL	60
DB	27	VFFQEXF--EDGWESRWKSEWKKNDENMAGEWNHTSGKNNGDAE-DKGIGTSEDYRFYAI	83
QY	61	SASFEPFSNKGQTLVVOFTVYKHEQNIDCGGGYVKLPNSLDQDTMKGDSYNIIMFGPDIC	120
DB	84	SAEYPEFSNKKDTLVLFQFSVKHEQKLDCCGGYVKLLGGVDQKKFGGDTSYSIMFGPDIC	143
QY	121	GPCTKKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSOVESGSL	179
DB	144	GYSTKKVHTLTLDGKNHLIKKDVPCETDQLTHVYTLILIRPDATYSILLIDNEEKQTGSI	202

Search completed: October 21, 2001, 03:07:44
Job time: 339 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:05:47 ; Search time 76 Seconds

(without alignments)

47.861 Million cell updates/sec

Title:

US-09-807-148-8

Perfect score:

318

Sequence:

1 GPGTKKHVIFNYKKNVLI.....PDNTVEVKIDNSQVSGSLE 60

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:

412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	318	100.0	60	21	AA192354
2	318	100.0	61	21	AA192352
3	318	100.0	180	21	AA192351
4	318	100.0	280	21	AA192355
5	318	100.0	400	21	AA192350
6	318	100.0	401	18	AA11156
7	318	100.0	417	10	AA192276
8	318	100.0	417	20	AA190927
9	318	100.0	417	21	AA192349
10	272	85.5	403	17	AA104171
11	258	81.1	49	21	AA192353
					Recombinant human
					Recombinant human
					Human vasostatin (
					Recombinant delta-
					Recombinant human
					Calreticulin. Hom
					60 kD Ro (Ro/SSA)
					Calreticulin. Hom
					Human MBP-calretic
					Flea calreticulin
					Recombinant human

12	242	76.1	336	12	AA12312	Partial sequence o
13	190	59.7	122	20	AA100924	Human cClqR binding
14	190	59.7	122	20	AA100926	Rat cClqR binding
15	182	57.2	122	20	AA100925	Mouse cClqR binding
16	177	55.7	415	22	AA166341	Castor bean calret
17	177	55.7	415	22	AA166343	Castor bean calret
18	175	55.0	385	21	AA132385	Human secreted pro
19	169	53.1	312	21	AA124809	Arabidopsis thalia
20	169	53.1	312	21	AA124809	Arabidopsis thalia
21	169	53.1	332	21	AA130998	Arabidopsis thalia
22	169	53.1	421	21	AA124608	Arabidopsis thalia
23	169	53.1	421	21	AA124608	Arabidopsis thalia
24	169	53.1	424	21	AA124607	Arabidopsis thalia
25	169	53.1	424	21	AA124607	Arabidopsis thalia
26	169	53.1	441	21	AA124607	Arabidopsis thalia
27	169	53.1	444	21	AA130996	Arabidopsis thalia
28	115	36.2	593	16	AA171094	Calnexin sequence.
29	100	31.4	417	21	AA177953	A. thaliana enviro
30	82	25.8	394	21	AA126285	Arabidopsis thalia
31	82	25.8	394	21	AA126285	Arabidopsis thalia
32	82	25.8	530	21	AA126284	Arabidopsis thalia
33	82	25.8	530	21	AA126284	Arabidopsis thalia
34	82	25.8	567	21	AA126284	Arabidopsis thalia
35	79	24.8	542	22	AA166342	Arabidopsis thalia
36	74	23.3	394	21	AA126285	Castor bean calnex
37	74	23.3	394	21	AA126285	Arabidopsis thalia
38	74	23.3	548	21	AA126285	Arabidopsis thalia
39	61.5	19.3	337	22	AA145695	P. falciparum yfg8
40	61	19.2	868	20	AA137731	Protein involved i
41	59.5	18.7	220	21	AA153217	Bridge-1 homologou
42	59	18.6	426	21	AA153217	Arabidopsis thalia
43	59	18.6	1844	21	AA18250	Plasmodium falcipa
44	58.5	18.4	1188	21	AA18183	Plasmodium falcipa
45	56.5	17.8	332	21	AA180287	Human G protein co

ALIGNMENTS

RESULT 1

AA192354

ID AA192354 standard; Protein; 60 AA.

XX AA192354;

XX AA192354;

DT 10-AUG-2000. (first entry)

XX Recombinant human calreticulin residues 121-180.

DE MBP-calreticulin; maltose binding protein; angio

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX endothe

XX Claim 4; Page 85; 99pp; English.
 XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 SQ Sequence 60 AA;

Query Match 100.0%; Score 318; DB 21; Length 60;
 Best Local Similarity 100.0%; Pred. No. 9.8e-38;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GPCTKKVHVIFNKGKLVNLIKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 60
 DB 1 9PPTKKVHVIFNKGKLVNLIKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 60

RESULT 2
 AAY92352
 ID AAY92352 standard; Protein; 61 AA.
 XX
 AC AAY92352;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Recombinant human calreticulin residues 120-180.
 XX
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytotatic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX
 OS Homo sapiens.
 OS Synthetic.

XX WO200020577-A1.
 XX 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US23240.
 XX
 PR 06-OCT-1998; 98US-0103438.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Tosato G, Pike SE, Yao L;
 XX
 DR WPI; 2000-303767/26.
 XX

XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 PT
 XX Claim 4; Page 82-83; 99pp; English.
 XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be

CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 SQ Sequence 61 AA;

Query Match 100.0%; Score 318; DB 21; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1e-37;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNKGKLVNLIKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 60
 DB 2 9PPTKKVHVIFNKGKLVNLIKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVSGSLE 61

RESULT 3
 AAY92351
 ID AAY92351 standard; Protein; 180 AA.
 XX
 AC AAY92351;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human vasostatin (calreticulin N-terminal 180 amino acids).
 XX
 KW MBP-calreticulin; maltose binding protein; vasostatin; N-terminal;
 KW angiogenesis; inhibition; endothelial cell; anti-angiogenic;
 KW neuroprotective; antidiabetic; cytotatic; dermalogical; hepatic;
 KW immunosuppressive; antiinflammatory; anti-atherosclerotic;
 KW gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.
 OS Synthetic.
 XX
 PN WO200020577-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US23240.
 XX
 PR 06-OCT-1998; 98US-0103438.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Tosato G, Pike SE, Yao L;
 XX
 DR WPI; 2000-303767/26.
 XX
 PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 XX
 PS Claim 4; Page 82; 99pp; English.

XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,

CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX Sequence 180 AA;
 SQ

Query Match 100.0%; Score 318; DB 21; Length 180;
 Best Local Similarity 100.0%; Pred. No. 4.1e-37;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKKHVIFNYGKNVLINKDKRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 Db 121 gpgtkkhvifnygknvlindkrckddefthlytlivrpndtyevkdnsvqesgsle 180

RESULT 4
 AAY92355
 ID AAY92355 standard; Protein: 280 AA.
 XX AC AAY92355;
 XX DT 10-AUG-2000 (first entry)
 XX DE Recombinant delta-120 calreticulin.
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; opthalmic.
 XX OS Homo sapiens.
 OS Synthetic.
 XX PN W0200020577-A1.
 XX PD 13-APR-2000.
 XX PF 05-OCT-1999; 99WO-US23240.
 XX PR 06-OCT-1998; 98US-0103438.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Tosato G, Pike SE, Yao L;
 XX DR WPI; 2000-303767/26.
 XX PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth
 PS Claim 4; Page 86; 99pp; English.
 XX This sequence comprises recombinant human calreticulin (AAY92355)
 CC missing the N-terminal 120 amino acids.
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The
 CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially

CC Kaposi's sarcoma (claimed).
 XX Sequence 280 AA;
 SQ

Query Match 100.0%; Score 318; DB 21; Length 280;
 Best Local Similarity 100.0%; Pred. No. 7.2e-37;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKKHVIFNYGKNVLINKDKRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 Db 1 gpgtkkhvifnygknvlindkrckddefthlytlivrpndtyevkdnsvqesgsle 60

RESULT 5
 AAY92350
 ID AAY92350 standard; Protein: 400 AA.
 XX AC AAY92350;
 XX DT 10-AUG-2000 (first entry)
 XX DE Recombinant human MBP-calreticulin.
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; opthalmic.
 XX OS Homo sapiens.
 XX PN W0200020577-A1.
 XX PD 13-APR-2000.
 XX PF 05-OCT-1999; 99WO-US23240.
 XX PR 06-OCT-1998; 98US-0103438.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Tosato G, Pike SE, Yao L;
 XX DR WPI; 2000-303767/26.
 XX PR N-ESDB; AAA09346, AAA09347.
 XX PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth
 XX PS Claim 4; Page 80-81; 99pp; English.
 XX Recombinant human MBP-calreticulin comprises the sequence of human
 CC calreticulin (see AAY92349) without the 17 N-terminal amino acids.
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The
 CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX Sequence 400 AA;
 SQ

Query Match 100.0%; Score 318; DB 21; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1.1e-36;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 |||
 DB 121 GPPTKKVHVFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
 |||

RESULT 6
 AAW11156
 ID AAW11156 standard; peptide; 401 AA.
 XX
 AC AAW11156;
 XX
 DT 31-MAY-1997 (first entry)
 XX
 DE Calreticulin.
 XX
 KW calreticulin; C-domain; restenosis; inhibitor.
 OS Homo sapiens.
 XX
 PN WO9636643-Al.
 XX
 PD 21-NOV-1996.
 XX
 PF 17-MAY-1996; 96WO-1B00471.
 XX
 PR 16-MAY-1996; 96US-0649417.
 PR 17-MAY-1995; 95US-0442844.
 XX
 PA (UYAL-) UNIV ALBERTA.
 XX
 PI Lucas A, Michalak M;
 XX
 WPI; 1997-012036/01.
 DR
 XX
 PT Inhibition of restenosis in patients - using calreticulin or a
 C-domain polypeptide of calreticulin or a variant with the same
 activity.
 XX
 PS Disclosure; Fig 1; 48pp; English.
 XX
 CC The present sequence is calreticulin. It and a C-domain derived peptide
 (AAW06736) are useful for treating a patient to inhibit restenosis. The
 CC calreticulin-type cpds. are administered either parenterally,
 CC intravenously or via a catheter and can target areas of vascular damage
 CC to inhibit or prevent restenosis.
 XX
 SQ Sequence 401 AA;

Query Match 100.0%; Score 318; DB 18; Length 401;
 Best Local Similarity 100.0%; Pred. No. 1.1e-36;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 |||
 DB 121 GPPTKKVHVFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
 |||

RESULT 7
 AAP92276
 ID AAP92276 standard; protein; 417 AA.
 XX
 AC AAP92276;
 XX
 DT 23-FEB-1990 (first entry)
 XX
 DE 60 kD Ro (Ro/SSA) antigen.
 XX
 WPI; 1999-180404/15.
 DR
 XX
 CC Sjorens syndrome; systemic lupus erythematosus.

XX Synthetic.
 OS
 XX
 PN WO8909273-A.
 XX
 PD 05-OCT-1989.
 XX
 PF 22-MAR-1989; 89WO-US01213.
 XX
 PR 22-MAR-1988; 88US-0171634.
 XX
 PA (TEXA) UNIV OF TEXAS SYST.
 XX
 PI Sontheimer RD, Capra JD, McCauliffe DP;
 XX
 DR WPI; 1989-309537/42.
 DR N-PSDB; AAP92276.
 XX
 PT DNA sequences encoding antigenic epitope(s) of Ro 60 kD autoantigen
 PT -used in immunoassays to detect rheumatic disease
 XX
 PS Disclosure; Fig 2; 88pp; English.
 XX
 CC Synthetic peptides corresp. to an epitopic core of Ro antigen are
 CC expressed recombinantly to detect autoantibodies, for identification
 CC of autoimmune diseases. These epitopes are AAs 24-36, 23-36, 188-209,
 CC or 241-255. The peptides may be substd. for ribonucleoprotein particle
 CC antigens.
 XX
 SQ Sequence 417 AA;

Query Match 100.0%; Score 318; DB 10; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.2e-36;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 |||
 DB 138 GPPTKKVHVFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 197
 |||

RESULT 8
 AAY00927
 ID AAY00927 standard; Protein; 417 AA.
 XX
 AC AAY00927;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE Calreticulin.
 XX
 KW Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
 KW CUB functional; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9907406-Al.
 XX
 PD 18-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-GB02430.
 XX
 PR 12-AUG-1997; 97GB-0016998.
 XX
 PA (UYLE-) UNIV LEICESTER.
 XX
 PI Schwaeble W;
 XX
 DR WPI; 1999-180404/15.
 XX

PT Use of a cClqR binding domain - to modulate complement ubiquitin
 XX (CUB) functionality.
 PS Disclosure; Page 26-27; 31pp; English.
 XX This sequence is calreticulin, a homologue of C1q and collectin receptor
 CC (cClqR). The invention relates to the use of a cClqR binding domain in a
 CC medicament to effect complement ubiquitin (CUB) functionality, and an
 CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 CC functionality. The cClqR binding domain, or its inhibitor, can be used to
 CC treat a human or animal body. Particularly an inhibitor is used to treat
 CC complement activation involved in the initiation and maintenance of
 CC inflammation, for example in myocardial infarction, brain ischaemia
 CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 CC erythematosus, burns, immune complex nephritis, and to treat amyloid
 CC plaques in Alzheimer's disease. The use of cClqR binding domain or
 CC inhibitor enables the CUB domain functionality to be modulated using a
 CC low molecular weight molecule.
 XX
 SQ Sequence 417 AA;

Query Match 100.0%; Score 318; DB 20; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.2e-36;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNKGKVLNKNKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 138 gpgtkkvhvifnkgkvnlinkdirckddefthlytlivrpdttyevkldnsqvesgsl 197

RESULT 9
 AAY92349
 ID AAY92349 standard; Protein: 417 AA.
 XX
 AC AAY92349;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human MBP-calreticulin.
 XX
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..17
 FT Protein 18
 FT /label= signal_peptide
 FT /label= mature_protein
 XX
 PN WO200020577-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US23240.
 XX
 PR 06-OCT-1998; 98US-0103438.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Tosato G, Pike SE, Yao L;
 XX
 DR WPI: 2000-303767/26.
 DR N-PSDB; AAA09346, AAA09347.
 XX
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 PT Disclosure; Page 79-80; 99pp; English.
 PS

XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 SQ Sequence 417 AA;

Query Match 100.0%; Score 318; DB 21; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.2e-36;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCTKKVHVIFNKGKVLNKNKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSLE 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 138 gpgtkkvhvifnkgkvnlinkdirckddefthlytlivrpdttyevkldnsqvesgsl 197

RESULT 10
 AAW04171
 ID AAW04171 standard; Protein: 403 AA.
 XX
 AC AAW04171;
 XX
 DT 12-DEC-1996 (first entry)
 XX
 DE Flea calreticulin PctCal403.
 XX
 KW Calreticulin; flea; haematophagous insect; allergic dermatitis;
 KW vaccine; therapy; PctCal403.
 XX
 OS Ctenocephalides felis.
 XX
 PN WO9628469-A1.
 XX
 PD 19-SEP-1996.
 XX
 PF 08-MAR-1996; 96WO-US03133.
 XX
 PR 09-MAR-1995; 95US-0401509.
 XX
 PA (HESK-) HESKA CORP.
 XX
 PI Rushlow KE, Stiegler GL;
 XX
 DR WPI: 1996-442861/44.
 DR N-PSDB; AAT39516;
 DR N-PSDB; AAT39517.
 XX
 XX Haematophagous insect calreticulin protein - used to reduce insect
 PT infestation and desensitise patients to allergic dermatitis
 PT Claim 5; Page 68-69; 86pp; English.
 PS
 PS
 CC Flea calreticulin protein PCTCall589 (AAW04171) is a calcium-binding
 CC protein found in the salivary glands of Ctenocephalides felis.
 CC Its amino acid sequence was deduced from a cDNA clone (AAT39516)
 CC obtd. from a salivary gland cDNA library. Recombinant PCTCall589
 CC can be produced in host cells transformed with a vector carrying
 CC calreticulin nucleic acids. Calreticulin alters the blood feeding
 CC behaviour of haematophagous insects and can be administered to an

AA	Sequence	49 AA;
SQ		

antigen is not subject to post-translational processing this suggests that most of the coding sequence is present. The three repeats are highly hydrophilic regions likely to be exposed on the surface of the antigen and highly immunogenic. Recombinant antigen expressed by the clone can be used stimulate T-cells of individuals infected by the parasite to proliferate and may be used as the basis for a vaccine against Onchocerciasis or river blindness.

XX.

XX

DT 28-MAY-1999 (first entry)
 XX Mouse cClqR binding domain protein sequence.
 DE Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
 XX CUB functionality; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy.
 XX Mus musculus.
 OS WO9907406-A1.
 XX 18-FEB-1999.
 XX 12-AUG-1998; 98WO-GB02430.
 XX 12-AUG-1997; 97GB-0016998.
 XX (UYLF-) UNIV LEICESTER.
 XX Schwaeble W;
 PI WPI; 1999-180404/15.
 DR N-PSDB; AAX27252.
 XX Use of a cClqR binding domain - to modulate complement ubiquitin
 (CUB) functionality.
 PS Claim 9; Page 24; 31pp; English.
 XX This sequence is a Clq and collectin receptor (cClqR) binding
 CC domain. The invention relates to the use of a cClqR binding domain in a
 CC medicament to effect complement ubiquitin (CUB) functionality, and an
 CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 CC functionality. The cClqR binding domain, or its inhibitor, can be used to
 CC treat a human or animal body. Particularly an inhibitor is used to treat
 CC complement activation involved in the initiation and maintenance of
 CC inflammation, for example in myocardial infarction, brain ischaemia
 CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 CC erythematosus, burns, immune complex nephritis, and to treat amyloid
 CC plaques in Alzheimer's disease. The use of cClqR binding domain or
 CC inhibitor enables the CUB domain functionality to be modulated using a
 CC low molecular weight molecule.
 XX Sequence 122 AA;

Query Match 57.2%; Score 182; DB 20; Length 122;
 Best Local Similarity 97.2%; Pred. No. 4.5e-18;
 Matches 35; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 25 RCKDDEFTHLVTLVPRDNTYEVKIDNSQVESGSLE 60
 DB 1 rckddeftlhlytlvprdqntyevkidnsqvesgsle 36

Search completed: October 21, 2001, 03:05:47
 Job time: 342 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:05:46 ; Search time 76 seconds
(without alignments)
143.583 Million cell updates/sec

Title: US-09-807-148-4
Perfect score: 971
Sequence: 1 EPAVYKQFLDGDGWTSRW.....PDNTVEVKIDNSQVSGSLE 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	971	100.0	180	AA192351	Human vasostatin (
2	971	100.0	400	AA192350	Recombinant human
3	971	100.0	417	AA192276	60 KD RO (Ro/SSA)
4	971	100.0	417	AA192349	Human MBP-calretic
5	966	99.5	417	AA1900927	Calreticulin. Hom
6	921	94.9	401	AA11156	Calreticulin. Hom
7	700	72.1	403	AA104171	Flea calreticulin
8	631	65.0	385	AA132385	Human secreted pro
9	541.5	55.8	336	AA112312	Partial sequence o
10	527	54.3	415	AA166341	Castor bean calret
11	527	54.3	415	AA166343	Castor bean calret

13	510	52.5	421	21	AA1924608	Arabidopsis thalia
14	510	52.5	421	21	AA1924607	Arabidopsis thalia
15	510	52.5	424	21	AA1924607	Arabidopsis thalia
16	510	52.5	441	21	AA1924607	Arabidopsis thalia
17	510	52.5	444	21	AA1924607	Arabidopsis thalia
18	327	33.7	61	21	AA1923352	Recombinant human
19	318	32.7	60	21	AA1923354	Recombinant human
20	318	32.7	280	21	AA1923355	Recombinant delta-
21	301.5	31.1	417	21	AA1923355	Calnexin sequence.
22	290	29.9	593	16	AA1923353	Recombinant human
23	258	26.6	49	21	AA1923353	Arabidopsis thalia
24	253	26.1	312	21	AA1923353	Arabidopsis thalia
25	253	26.1	312	21	AA1923353	Arabidopsis thalia
26	253	26.1	332	21	AA1923353	Arabidopsis thalia
27	250.5	25.8	84	21	AA1923353	Arabidopsis thalia
28	238.5	24.6	530	21	AA1923353	Arabidopsis thalia
29	238.5	24.6	530	21	AA1923353	Arabidopsis thalia
30	238.5	24.6	567	21	AA1923353	Arabidopsis thalia
31	227.5	23.4	542	22	AA1923353	Castor bean calnex
32	200.5	20.6	532	21	AA1923353	Arabidopsis thalia
33	200.5	20.6	548	21	AA1923353	Arabidopsis thalia
34	190	19.6	122	20	AA1923353	Human cClqR bindin
35	190	19.6	122	20	AA1923353	Rat cClqR binding
36	182	18.7	122	20	AA1923353	Mouse cClqR bindin
37	145	14.9	91	21	AA1923353	Arabidopsis thalia
38	145	14.9	99	21	AA1923353	Arabidopsis thalia
39	124.5	12.8	221	21	AA1923353	Human prostate can
40	114.5	11.8	162	21	AA1923353	Human secreted pro
41	114.5	11.8	162	21	AA1923353	Human secreted pro
42	112.5	11.6	394	21	AA1923353	Arabidopsis thalia
43	112.5	11.6	394	21	AA1923353	Arabidopsis thalia
44	110.5	11.4	125	21	AA1923353	Arabidopsis thalia
45	110.5	11.4	126	21	AA1923353	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AA192351

ID AA192351 standard; Protein: 180 AA.

AC AA192351;

DT 10-AUG-2000 (first entry)

DE Human vasostatin (calreticulin N-terminal 180 amino acids).

XX MBP-calreticulin; maltose binding protein; vasostatin; N-terminal;
 KW angiogenesis; inhibition; endothelial cell; anti-angiogenic;
 KW neuroprotective; antidiabetic; cytostatic; dermatologic; hepatic;
 KW immunosuppressive; antiinflammatory; anti-atherosclerotic;
 KW gastrointestinal; anti-arthritis; ophthalmic.

OS Homo sapiens.

OS Synthetic.

PN WO200020577-A1.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.

XX 05-OCT-1998; 98US-0103438.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Torato G, Pike SE, Yao L;

XX WPL; 2000-303767/26.

PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,

PT useful for suppressing tumor growth
 PS Claim 4; Page 82; 99pp; English.
 XX
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 SQ Sequence 180 AA;

Query Match 100.0%; Score 971; DB 21; Length 180;
 Best Local Similarity 100.0%; Pred. No. 2.4e-99;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 60
 DB 1 EPAYVKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 60
 QY 61 SASFEPFSNKGQTLVVQFTVKHEQNDICGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120
 DB 61 SASFEPFSNKGQTLVVQFTVKHEQNDICGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120
 QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 180
 DB 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 180

RESULT 2
 AAY92350
 ID AAY92350 standard; Protein: 400 AA.
 AC AAY92350;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Recombinant human MBP-calreticulin.
 XX
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatological; immunosuppressive; anti-inflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX
 OS Homo sapiens.
 XX
 PN WO200020577-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US23240.
 XX
 PR 06-OCT-1998; 98US-0103438.
 XX
 PA (US) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Tosato G, Pike SE, Yao L;
 XX
 DR WPI: 2000-303767/26.
 DR N-PSDB; AAA09346, AAA09347.
 XX

PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 XX
 PS Claim 4; Page 80-81; 99pp; English.
 XX
 CC Recombinant human MBP-calreticulin comprises the sequence of human
 CC calreticulin (see AAY92349) without the 17 N-terminal amino acids.
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The
 CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 SQ Sequence 400 AA;

Query Match 100.0%; Score 971; DB 21; Length 400;
 Best Local Similarity 100.0%; Pred. No. 7.3e-99;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 60
 DB 1 EPAYVKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFYAL 60
 QY 61 SASFEPFSNKGQTLVVQFTVKHEQNDICGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120
 DB 61 SASFEPFSNKGQTLVVQFTVKHEQNDICGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120
 QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 180
 DB 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 180

RESULT 3
 AAP92276
 ID AAP92276 standard; protein: 417 AA.
 XX
 AC AAP92276;
 XX
 DT 23-FEB-1990 (first entry)
 XX
 DE 60 kD Ro (Ro/SSA) antigen.
 XX
 KW Sjorens syndrome; systemic lupus erythematosus.
 XX
 OS Synthetic.
 XX
 PN WO8909273-A.
 XX
 PD 05-OCT-1989.
 XX
 PF 22-MAR-1989; 89WO-US01213.
 XX
 PR 22-MAR-1988; 88US-0171634.
 XX
 PA (TEXA) UNIV OF TEXAS SYST.
 XX
 PI Sontheimer RD, Capra JD, McCauliffe DP;
 XX
 DR WPI: 1989-309537/42.
 DR N-PSDB; AAP92276.
 XX

PT DNA sequences encoding antigenic epitope(s) of Ro 60 kD autoantigen
 PT - used in immunoassays to detect rheumatic disease
 PS Disclosure: Fig 2; 88pp; English.
 CC Synthetic peptides corresp. to an epitopic core of Ro antigen are
 CC expressed recombinantly to detect autoantibodies for identification
 CC of autoimmune diseases. These epitopes are AAs 24-36, 23-36, 188-209,
 CC or 241-255. The peptides may be substd. for ribonucleoprotein particle
 CC antigens.
 XX Sequence 417 AA;
 SQ

Query Match 100.0%; Score 971; DB 10; Length 417;
 Best Local Similarity 100.0%; Pred. No. 7.8e-99;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVFKQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFVAL 60
 Db 18 epavyfkeqfldggwtsrwieskhksdfgkfvlsagkygdeekdkglqtsqdarfyal 77
 QY 61 SASPEPFSNKGQTLVVOFTVKHEQNIDCGGYVKLFPSNLSLQDTMDHGDSEYNIMFGPDIC 120
 Db 78 sasfepfsnkgqtlvvqftvkheqndcggyvklfpsnldqtdmhdgseynimfgpdic 137
 QY 121 GPGTKKHVHIFNYKGNVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
 Db 138 gpgtkkhvifnykgnvinkdirckddefthlytlivrpndtyevkidnsqvesgsle 197

RESULT 4
 AAY92349
 ID AAY92349 standard; Protein: 417 AA.
 XX
 AC AAY92349;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human MBP-calreticulin.
 XX
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytotatic; dermatological; immunosuppressive; antinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..17
 FT Protein /label= signal_peptide
 FT /label= mature_protein
 XX
 PN WO200020577-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US23240.
 XX
 PR 06-OCT-1998; 98US-0103438.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Tosato G, Pike SE, Yao L;
 XX
 DR WPI; 2000-303767/26.
 DR N-PSDB; AAA09346, AAA09347.
 XX
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth
 PT
 PS Disclosure: Page 79-80; 99pp; English.

XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX Sequence 417 AA;
 SQ

Query Match 100.0%; Score 971; DB 21; Length 417;
 Best Local Similarity 100.0%; Pred. No. 7.8e-99;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVFKQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSQDARFVAL 60
 Db 18 epavyfkeqfldggwtsrwieskhksdfgkfvlsagkygdeekdkglqtsqdarfyal 77
 QY 61 SASPEPFSNKGQTLVVOFTVKHEQNIDCGGYVKLFPSNLSLQDTMDHGDSEYNIMFGPDIC 120
 Db 78 sasfepfsnkgqtlvvqftvkheqndcggyvklfpsnldqtdmhdgseynimfgpdic 137
 QY 121 GPGTKKHVHIFNYKGNVINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
 Db 138 gpgtkkhvifnykgnvinkdirckddefthlytlivrpndtyevkidnsqvesgsle 197

RESULT 5
 AAY00927
 ID AAY00927 standard; Protein: 417 AA.
 XX
 AC AAY00927;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE Calreticulin.
 XX
 KW C1q and collectin receptor; cclqr binding domain; complement ubiquitin;
 KW C1q functionality; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9907406-A1.
 XX
 PD 18-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-GB02430.
 XX
 PR 12-AUG-1997; 97GB-0016998.
 XX
 PA (OYLE-) UNIV LEICESTER.
 XX
 PI Schwaebler W;
 XX
 DR WPI; 1999-180404/15.
 XX
 PT Use of a cclqr binding domain - to modulate complement ubiquitin
 PT (CUB) functionality.
 XX

PS Disclosure; Page 26-27; 31pp; English.

XX This sequence is calreticulin, a homologue of C1q and collectin receptor (C1qR). The invention relates to the use of a C1qR binding domain in a CC medicament to effect complement ubiquitin (CUB) functionality, and an CC inhibitor of the C1qR binding domain in a medicament to inhibit CUB CC functionality. The C1qR binding domain, or its inhibitor, can be used to CC treat a human or animal body. Particularly an inhibitor is used to treat CC complement activation involved in the initiation and maintenance of CC inflammation, for example in myocardial infarction, brain ischaemia CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus CC erythematosus, burns, immune complex nephritis, and to treat amyloid CC plaques in Alzheimer's disease. The use of C1qR binding domain or CC inhibitor enables the CUB domain functionality to be modulated using a CC low molecular weight molecule.

XX Sequence 417 AA;

Query Match 99.5%; Score 966; DB 20; Length 417;
Best Local Similarity 99.4%; Pred. No. 2.8e-98;
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 E PAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60
DB 18 E PAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77
QY 61 SASFEPFSNKGQTLVQVFTVKHEQNIDCGGYYVKLPNSLDQTMHGDSEYINMFGPDIC 120
DB 78 SASFEPFSNKGQTLVQVFTVKHEQNIDCGGYYVKLPNSLDQTMHGDSEYINMFGPDIC 137
QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
DB 138 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 197

RESULT 6
AAW041156
ID AAW041156 standard; peptide; 401 AA.
AC AAW041156;
DT 31-MAY-1997 (first entry)
XX Calreticulin.
DE Calreticulin; C-domain; restenosis; inhibitor.
KW Homo sapiens.
OS
XX WO9636643-Al.
PN 21-NOV-1996.
PD 17-MAY-1996; 96WO-IB00471.
PF 16-MAY-1996; 96US-0649417.
PR 17-MAY-1995; 95US-0442844.
XX (UYAL-) UNIV ALBERTA.
PA Lucas A, Michalak M;
PI WPI; 1997-012036/01.
XX Inhibition of restenosis in patients - using calreticulin or a
PT C-domain polypeptide of calreticulin or a variant with the same.
PT activity.
XX
PS Disclosure; Fig 1; 48pp; English.
XX The present sequence is calreticulin. It and a C-domain derived peptide
CC (AAW06736) are useful for treating a patient to inhibit restenosis. The

CC calreticulin-type cpds. are administered either parenterally,
CC intravenously or via a catheter and can target areas of vascular damage
XX to inhibit or prevent restenosis.
SQ Sequence 401 AA;

Query Match 94.9%; Score 921; DB 18; Length 401;
Best Local Similarity 95.0%; Pred. No. 2.4e-93;
Matches 171; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 E PAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60
DB 1 E PAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60
QY 61 SASFEPFSNKGQTLVQVFTVKHEQNIDCGGYYVKLPNSLDQTMHGDSEYINMFGPDIC 120
DB 61 SARFEPFSNKGQTLVQVFTVKHEQNIDCGGYYVKLPNSLDQTMHGDSEYINMFGPDIC 120
QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
DB 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180

RESULT 7
AAW041171
ID AAW041171 standard; Protein; 403 AA.
AC AAW041171;
XX 12-DEC-1996 (first entry)
DT
XX Flea calreticulin PctCal403.
DE
XX Calreticulin; flea; haematophagous insect; allergic dermatitis;
KW vaccine; therapy; PctCal403.
XX Ctenocephalides felis.
OS
XX WO9628469-Al.
PN 19-SEP-1996.
PD 08-MAR-1996; 96WO-US03133.
PF 09-MAR-1995; 95US-0401509.
PR (HESK-) HESKA CORP.
XX
PA Rushlow KE, Stiegler GL;
PI WPI; 1996-442861/44.
DR N-PSDB; AAT39516;
DR N-PSDB; AAT39517.
XX
PT Haematophagous insect calreticulin protein - used to reduce insect
PT infestation and desensitise patients to allergic dermatitis
XX
PS Claim 5; Page 68-69; 86pp; English.
XX Flea calreticulin protein PctCall589 (AAW041171) is a calcium-binding
CC protein found in the salivary glands of ctenocephalides felis.
CC Its amino acid sequence was deduced from a cDNA clone (AAT39516)
CC obtd. from a salivary gland cDNA library. Recombinant PctCall589
CC can be produced in host cells transformed with a vector carrying
CC calreticulin nucleic acids. Calreticulin alters the blood feeding
CC behaviour of haematophagous insects and can be administered to an
CC animal to reduce infestation. It reduces calreticulin activity in
CC insects, so reducing the insect burden on an animal. Calreticulin
CC can be used to elicit an immune response, thereby desensitising an
CC animal to allergic dermatitis caused by fleas, mosquitoes or
CC Culicoides.

SQ Sequence 403 AA;

Query Match 72.1%; Score 700; DB 17; Length 403;
Best Local Similarity 71.7%; Pred. No. 6.1e-69;
Matches 129; Conservative 23; Mismatches 26; Indels 2; Gaps 2;

QY 2 PAVYFKQFLDGGWTSRWIESKHK-SDFGKFLVSSGKFGYDEKDKGLOTSQDARFVAL 60
Db 19 pveleenfvd-dwtntwyshepbkefghvhtagkfyndaeakgltqgqdarifal 77
QY 61 SASPEFNSKQTLVVOFTVKHEQNIIDCGGYVKLFPSNLSLDTQDMHGDSSEYNIMFGPDIC 120
Db 78 shkfkpsnkaktlvqfsvkheqniidcggykkgfsvnqkdmhgespyeimfgpdic 137
QY 121 GPGTKKHVIFNYGKKNVINKDRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
Db 138 dpgtkvhvifsygkknvinkdrckddvthvtytlvvpkdpntyevlidnekvesgule 197

RESULT 8
AAB32385
ID AAB32385 standard; Protein: 385 AA.
XX
AC AAB32385;
DT 16-JAN-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 15 SEQ ID NO:71.
XX
KW Human; secreted protein; cytostatic; immunostimulant; antiproliferative;
KW cardiant; antiarrhythmic; antiviral; antibacterial; antifungal; cancer;
KW antiparasitic; neuroprotective; nootropic; antiinflammatory; anti-HIV;
KW antiangiogenic; antiarteriosclerotic; diagnosis; immune disorder; AIDS;
KW autoimmune disease; haematopoietic cell disorder; blood protein disorder;
KW agammaglobulinaemia; hyperproliferative disease; Gaucher's disease;
KW cardiovascular disorder; congenital heart defect; pulmonary atresia;
KW arrhythmia; ischaemia; angiogenesis related disorder; Crohn's disease;
KW atherosclerosis; neurological disease; Alzheimer's disease;
KW Huntington's; infectious disease; cat-scratch disease.
XX
OS Homo sapiens.
XX
PN WO200047602-A1.
XX
PD 17-AUG-2000.
XX
PF 08-FEB-2000; 2000WO-US03062.
XX
PR 10-FEB-1999; 99US-0119468.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Soppet DR, Moore PA;
PI Shi Y, Lafleur DW, Olsen HS, Florence KA, Komatsoulis G;
XX
XX WPI: 2000-543578/49.
DR N-PSDB; AAC55204.
XX
XX New human nucleic acids encoding secreted proteins, useful in the
PT treatment, prevention or diagnosis of immune disorders (e.g. autoimmune
PT diseases), blood protein disorders and hyperproliferative diseases
PT (e.g. Gaucher's disease) -
XX
XX Claim 11; Page 434-435; 488pp; English.
XX
CC The polynucleotide sequences given in AAC55190 to AAC55235 encode the
CC human secreted proteins given in AAB32371 to AAB32484. Human secreted
CC proteins have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic;
CC immunostimulant; antiproliferative; cardiant; antiarrhythmic; antiviral;
CC antibacterial; antifungal; antiparasitic; neuroprotective; nootropic;
CC antiinflammatory; antiangiogenic; anti-HIV; and antiarteriosclerotic. The

CC polynucleotides and polypeptides, or their agonists and antagonists, can
CC be used for treating, preventing or diagnosing immune disorders (e.g.
CC cancer, autoimmune diseases), disorders of haematopoietic cells, blood
CC protein disorders (e.g. agammaglobulinaemia), hyperproliferative
CC diseases (e.g. Gaucher's disease), cardiovascular disorders (e.g.
CC congenital heart defects, pulmonary atresia, arrhythmias, ischaemia),
CC angiogenesis related disorders (e.g. Crohn's disease, atherosclerosis),
CC neurological diseases (e.g. Alzheimer's disease, Huntington's chorea),
CC infectious diseases (e.g. AIDS, cat-scratch disease and other bacterial,
CC viral, parasitic or fungal diseases). AAC55181 to AAC55189 and AAB32370
CC represent sequences used in the exemplification of the present
XX invention.
XX
SQ Sequence 385 AA;
Query Match 65.0%; Score 631; DB 21; Length 385;
Best Local Similarity 64.4%; Pred. No. 2.4e-61;
Matches 114; Conservative 30; Mismatches 33; Indels 0; Gaps 0;

QY 4 VYFKQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDEKDKGLOTSQDARFVALSAS 63
Db 21 vyfgeefldgehrnrlwqstndrsrghfrlssgkfyghkekkgltqgrfyaiair 80
QY 64 FEPSNKGOTLVVOFTVKHEQNIIDCGGYVKLFPSNLSLDTQDMHGDSSEYNIMFGPDICGP 123
Db 81 fkpfsnkgotlvqvkvkheqniidcggykvpadiqknlgksqyymfgpdicgfd 140
QY 124 TKKHVIFNYGKKNVINKDRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 180
Db 141 ikkhvhlhfnknyhenkklirckvgtghlytlilrpdlsydvkldgsgiesgslie 197

RESULT 9
AAR12312
ID AAR12312 standard; Protein: 336 AA.
XX
AC AAR12312;
DT 29-AUG-1991 (first entry)
XX
DE Partial sequence of Onchocera volvulus 42 kD antigen.
XX
KW River blindness; onchocerciasis; vaccine; antigen; parasite.
XX
OS Onchocerca volvulus.
XX
XX Key Location/Qualifiers
XX Region 160..166
XX /label= repeat unit
XX /note= "hydrophilic"
XX Region 177..183
XX /label= repeat unit
XX /note= "hydrophilic"
XX Region 195..201
XX /label= repeat unit
XX /note= "hydrophilic"
XX
XX US5021342-A.
XX
XX 04-JUN-1991.
XX
XX 30-JUN-1988; 88US-0214264.
XX
XX 30-JUN-1988; 88US-0214264.
XX
XX (UYHO-) UNIV HOSPITALS CLEV.
XX
XX Greene BM, Unnasch TR;
XX WPI: 1991-185179/25.
XX N-PSDB; AAQ11987.
XX

PT DNA encoding Onchocerca volvulus antigen - used to express
PT recombinant antigen for vaccine against onchocerciasis or river
PT blindness.

PS Disclosure; Fig 7; 20pp; English.

XX The sequence was deduced from a cDNA clone lambda RAL-1 prepd.
XX from RNA isolated from nodules excised from patients infected
XX with O. volvulus. The N-terminal is incomplete, however Abs
XX which specifically bind to protein prepd. from induced cultures
XX of lambda RAL-1 lysogens recognise a single polypeptide of mol.
XX wt. 42,000 in extracts of adult worms. Analysis deduced sequence
XX suggests that it encodes a protein of mol. wt. 39,130. If the
XX antigen is not subject to post-translational processing this
XX suggests that most of the coding sequence is present. The three
XX repeats are highly hydrophilic regions likely to be exposed on
XX the surface of the antigen and highly immunogenic. Recombinant
XX antigen expressed by the clone can be used stimulate T-cells of
XX individuals infected by the parasite to proliferate and may be
XX used as the basis for a vaccine against Onchocerciasis or river
XX blindness.

SQ Sequence 336 AA;

Query Match 55.8%; Score 541.5; DB 12; Length 336;
Best Local Similarity 68.5%; Pred. No. 1.5e-51;
Matches 98; Conservative 21; Mismatches 23; Indels 1; Gaps 1;

QY 39 FYGDEKDGLOTSQDAREYALSASFE-PFSNKGOTLVVQFTVKHEQNDCCGGYKLPF 97
DB 1 fyydvdvdkgkktqtdakfysigaktdksfsnkgkslvqfsvkheqelcgggyvklma 60
QY 98 NSLDQDTHMGDSEYNTMFGPDICGPGTKKVHVIFNYKGNVLINKDIRCKDDFTHLTYL 157
DB 61 sdvnledshgetpyhmfmpdpcgptkvhvifnykgnhmikkdirckddvftllytl 120
QY 158 IVRPDNTYEVKIDNSQVSGSL 180
DB 121 invsdntyevidgekaesgele 143

RESULT 10
AAB66341
ID AAB66341 standard; Protein; 415 AA.

AC AAB66341;

DT 05-APR-2001 (first entry)

DE Castor bean calreticulin SEQ ID NO: 2.

XX Castor bean; calreticulin; calcium binding; er; endoplasmic reticulum;
XX calnexin; promoter; resistance.

OS Ricinus communis.

PN US6171864-B1.

PD 09-JAN-2001.

PF 05-JUL-1996; 96US-0675816.

PR 05-JUL-1996; 96US-0675816.

PA (PION-) PIONEER HI-BRED INT INC.

PI Coughlan SJ, Winfrey RJ;

XX WPI; 2001-122335/13.

DR N-PSDB; AAF29741, AAF29742.

PT New nucleic acid molecules encoding a calcium binding chaperone protein

PT in endoplasmic reticulum, calreticulin, and calreticulin promoter
PT sequences, useful for producing foreign gene products in plant cells
PS Claim 1; Fig 3; 45pp; English.

XX The present invention provides the protein and coding sequences for the
XX castor bean calreticulin protein and the calreticulin promoter sequence.
XX In addition, the castor bean calnexin protein, coding sequence and
XX promoter are also described. Calreticulin and calnexin are calcium
XX binding proteins found in the endoplasmic reticulum (er). The
XX calreticulin promoter sequence can be used in vectors to promote the
XX expression of foreign genes, particularly resistance genes, in plant
XX cells.

SQ Sequence 415 AA;

Query Match 54.3%; Score 527; DB 22; Length 415;
Best Local Similarity 55.3%; Pred. No. 7.9e-50;
Matches 99; Conservative 31; Mismatches 43; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHKSD---FGKFLVLSGKFYGDSEKDKGLOTSQDAREYAL 60
DB 22 vffeerf--edgwenrwksdwkdkentagewnytsqkwngd-pndkgtdsedyrfyal 78
QY 61 SASFEPFSNKGOTLVVQFTVKHEQNDCCGGYKLPFNSLDQDTHMGDSEYNTMFGPDIC 120
DB 79 saefsfnsnkdktlvqfsvkheqkldcggymkllssstdqkkgdtpyslmfgpdlc 138
QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLTYLIVRPDNTYEVKIDNSQVSGSL 179
DB 139 gystkvvhalnyndtnhlikvcpetdqlthvtytlvirdatysilidnvektgsl 197

RESULT 11
AAB66343
ID AAB66343 standard; Protein; 415 AA.

AC AAB66343;

DT 05-APR-2001 (first entry)

DE Castor bean calreticulin.

XX Castor bean; calreticulin; calcium binding; er; endoplasmic reticulum;
XX calnexin; promoter; resistance.

OS Ricinus communis.

PN US6171864-B1.

PD 09-JAN-2001.

PF 05-JUL-1996; 96US-0675816.

PR 05-JUL-1996; 96US-0675816.

PA (PION-) PIONEER HI-BRED INT INC.

PI Coughlan SJ, Winfrey RJ;

XX WPI; 2001-122335/13.

DR N-PSDB; AAF29755.

PT New nucleic acid molecules encoding a calcium binding chaperone protein
PT in endoplasmic reticulum, calreticulin, and calreticulin promoter
PT sequences, useful for producing foreign gene products in plant cells
PS Disclosure; Fig 1B; 45pp; English.

XX The present invention provides the protein and coding sequences for the
XX castor bean calreticulin protein and the calreticulin promoter sequence.
XX In addition, the castor bean calnexin protein, coding sequence and

CC promoter are also described. Calreticulin and calnexin are calcium
 CC binding proteins found in the endoplasmic reticulum (er). The
 CC calreticulin promoter sequence can be used in vectors to promote the
 CC expression of foreign genes, particularly resistance genes, in plant
 CC cells.
 XX
 SQ Sequence 415 AA;

Query Match 54.3%; Score 527; DB 22; Length 415;
 Best Local Similarity 55.3%; Pred. No. 7.9e-50;
 Matches 99; Conservative 31; Mismatches 43; Indels 6; Gaps 3;
 QY 4 VYKQFLDGGWTSRWIESKHSD---FGKFLSSGKFGDEKDKLQTSQDARFVAL 60
 Db 22 vffeerf--edgwenrwwksdkkidentagewnytskgwngd-pndkgigtseyrfyai 78
 QY 61 SASPEPFSKGTQTLVQFTVRHEQNIDCGGKLVFPNSLQDTDMHGDSKYNIMFGPDIC 120
 Db 79 saefpfsnkgtktlvfqsfsvheqkldcggymkllssstqdkkfsgdtpysimfgpdic 138
 QY 121 GPGTKKHVIFNYGKKNVLINIKRCKDDETHLYTLIVRPDNTVEYKIDNSQVESGSL 179
 Db 139 gystkkvhailnyndtnhlikvepcetdqtthvtylvirpdatysilidhvektqtsl 197

RESULT 12
 AAG24608
 ID AAG24608 standard; Protein: 421 AA.
 XX
 AC AAG24608;
 XX
 DT 17-OCN-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 28350.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.
 XX EPI033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145195.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145095.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 19-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.

```
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      52.5%; Score 510; DB 21; Length 421;
Best Local Similarity 53.9%; Pred. No. 6.le-48;
Matches 97; Conservative 30; Mismatches 47; Indels 6; Gaps 3;

QY 3 AVFEKQFLDGDWTSRWTESKHSD---FGKFLSSGRFYGDDEKDKGLQTSQDAREYA 59
   ||:|:| ||| :|:|:|:| | | :|:|:|:| |||
Db 20 avifeerf--ddgwenrwksekwdndtagewkhtagnwgd-andkqigtseyrfya 76

QY 60 LSASFEPFSNKQTLVQPTVRKHEQNIQDCGGYVKLFNPSLDQTDHMGDSEYNIMEGPDI 119
   :|| | |||| :||| :||| |||| :||| |||| :||| ||||
Db 77 isaefpfsnkktlvfqfsvkheqklcggymkllsgdvdkkfggdpysimfgpdi 136

QY 120 CGPCHKKVHVIFNKGKVNLIKNDIRCKDDETHLYTLVRPDNTVEYKIDNSQVSESSL 179
   || |||| | | | | | | | | | | | | | | | | | | | | | |
Db 137 cdystkvhailtyneanhlikdvpcetdqthvyflrpdatsyllidnvektgsl 196

RESULT 14
AAG24607
ID AAG24607 standard; Protein; 424 AA.
AC AAG24607;
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28349.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
```


PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

```
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 52.5%; Score 510; DB 21; Length 424;
Best Local Similarity 53.9%; Pred. NO. 6.1e-48;
Matches 97; Conservative 30; Mismatches 47; Indels 6; Gaps 3;

QY 3 AVYFKEQFLDGDGWTNRWIESKHKSD---FCKFVLSGKFYGDDEKDKGLQTSQDARFYA 59
   |||:| ||| :||:| | | :|: || |||:| |||
Db 23 avifeerf--ddgwenrwwksekdkddntagewkhtagnwsgd-andkgigtsedyrfya 79

QY 60 LSASFEPFSNKGQTLVVOFTVKHEQNIDCGGYVKLFPPNSLDQDTMHGDSSEYNIMFGPDI 119
   :|| | |||| :||| :|||:| |||||:| | | :||: |||:| |||
Db 80 isaepefsnkdktlvfsvkheqldcggymkllsgdvdkkfggdtypsimgpdi 139

QY 120 CGPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYTLVVRPONTYEVKIDNSQVESGSL 179
   || ||||| | | | |||: | | :|||:| ||| | | :||| :|||
Db 140 cgystkkvhaltyneanhlikkdvpcetdqlthvyffllrpdatsilldnvekqtgs1 199
```

Search completed: October 21, 2001, 03:05:46
Job time: 341 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:05:47 ; Search time 76 Seconds
(without alignments)
39.087 Million cell updates/sec

Title: US-09-807-148-6
Perfect score: 258
Sequence: 1 NYGKNVINKIRKDEF.....PDNTYEVKIDNSQVESGSL 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_0601.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	258	100.0	49	21	AA192353
2	258	100.0	60	21	AA192354
3	258	100.0	61	21	AA192352
4	258	100.0	180	21	AA192351
5	258	100.0	280	21	AA192355
6	258	100.0	400	21	AA192350
7	258	100.0	401	18	AA192356
8	258	100.0	417	10	AA192357
9	258	100.0	417	20	AA192358
10	258	100.0	417	21	AA192359
11	218	84.5	403	17	AA192360

12	190	73.6	122	20	AA192354	Human cClqR bindin
13	190	73.6	122	20	AA192354	Rat cClqR binding
14	182	70.5	122	20	AA192354	Mouse cClqR binding
15	182	70.5	122	20	AA192354	Partial sequence o
16	145	56.2	385	21	AA192355	Human secreted pro
17	143	55.4	415	22	AA192356	Castor bean calret
18	143	55.4	415	22	AA192356	Castor bean calret
19	135	52.3	312	21	AA192357	Arabidopsis thalia
20	135	52.3	312	21	AA192357	Arabidopsis thalia
21	135	52.3	332	21	AA192358	Arabidopsis thalia
22	135	52.3	421	21	AA192359	Arabidopsis thalia
23	135	52.3	421	21	AA192359	Arabidopsis thalia
24	135	52.3	424	21	AA192360	Arabidopsis thalia
25	135	52.3	424	21	AA192360	Arabidopsis thalia
26	135	52.3	441	21	AA192361	Arabidopsis thalia
27	135	52.3	444	21	AA192362	Arabidopsis thalia
28	100	38.8	593	16	AA192363	Calnexin sequence.
29	97.5	37.8	417	21	AA192364	A. thaliana enviro
30	68	26.4	394	21	AA192365	Arabidopsis thalia
31	68	26.4	394	21	AA192365	Arabidopsis thalia
32	68	26.4	530	21	AA192366	Arabidopsis thalia
33	68	26.4	530	21	AA192366	Arabidopsis thalia
34	68	26.4	542	22	AA192367	Castor bean calnex
35	68	26.4	567	21	AA192368	Arabidopsis thalia
36	65	25.2	394	21	AA192369	Arabidopsis thalia
37	65	25.2	532	21	AA192370	Arabidopsis thalia
38	65	25.2	548	21	AA192371	Arabidopsis thalia
39	61.5	23.8	337	22	AA192372	P. falciparum yfpg
40	57.5	22.3	1188	21	AA192373	Plasmodium falcipa
41	56.5	21.9	823	21	AA192374	Arabidopsis thalia
42	55	21.3	1844	21	AA192375	Plasmodium falcipa
43	54	20.9	405	15	AA192376	Human corticostero
44	54	20.9	405	15	AA192376	Human varient cort
45	54	20.9	405	15	AA192376	Corticosteroid bin

ALIGNMENTS

RESULT. 1
AA192353
ID AA192353 standard; Protein; 49 AA.
XX
AC AA192353;
XX
DT 10-AUG-2000 (first entry)
XX
DE Recombinant human calreticulin residues 132-180.
XX
KW MBP-calreticulin; maltose binding protein; angioinogenesis; inhibition;
KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
KW cytotstatic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200020577-A1.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US23240.
XX
PR 06-OCT-1998; 98US-0103438.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Tosato G, Pike SE, Yao L;
XX Waf; 2000-303767/26.
XX
DR Inhibiting endothelial cell growth and angiogenesis using calreticulin,
XX useful for suppressing tumor growth
PT

XX PS Claim 4; Page 82-83; 99pp; English.

XX CC A novel method of inhibiting endothelial cell growth comprises

CC contacting the cells with calreticulin (or its fragments/variants).

CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,

CC tumor growth and/or endothelial cell growth (claimed). The method may be

CC used for inhibiting angiogenesis in a patient. The angiogenesis is

CC associated with a disease other than a tumor that is associated with

CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,

CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune

CC inflammation, atherosclerosis, excessive wound repair, retinal

CC neovascularization, macular degeneration, corneal graft rejection,

CC contact lens overwear, Crohn's disease, non-immune inflammation,

CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,

CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's

CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may

CC also be used for treating/inhibiting tumor growth especially

CC Kaposi's sarcoma (claimed).

XX SQ Sequence 49 AA;

Query Match 100.0%; Score 258; DB 21; Length 49;

Best Local Similarity 100.0%; Pred. No. 1e-31;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGNVLINKDIRCKDDEFTHLTVLRPNTYEVKIDNSQVSGSLE 49

DB 1 NYKGNVLINKDIRCKDDEFTHLTVLRPNTYEVKIDNSQVSGSLE 49

RESULT 2

AAAY92354

ID AAAY92354 standard; Protein; 60 AA.

XX AC AAAY92354;

XX DT 10-AUG-2000 (first entry)

DE Recombinant human calreticulin residues 121-180.

XX MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;

KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;

KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;

KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.

OS Synthetic.

XX WO200020577-A1.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.

XX 06-OCT-1998; 98US-0103438.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Tosato G, Pike SE, Yao L;

XX WPI; 2000-303767/26.

XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,

XX useful for suppressing tumor growth

XX Claim 4; Page 85; 99pp; English.

XX A novel method of inhibiting endothelial cell growth comprises

XX contacting the cells with calreticulin (or its fragments/variants).

XX Fragments of calreticulin causes at least 40% inhibition of angiogenesis,

XX tumor growth and/or endothelial cell growth (claimed). The method may be

CC used for inhibiting angiogenesis in a patient. The angiogenesis is

CC associated with a disease other than a tumor that is associated with

CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,

CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune

CC inflammation, atherosclerosis, excessive wound repair, retinal

CC neovascularization, macular degeneration, corneal graft rejection,

CC contact lens overwear, Crohn's disease, non-immune inflammation,

CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,

CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's

CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may

CC also be used for treating/inhibiting tumor growth especially

CC Kaposi's sarcoma (claimed).

XX SQ Sequence 60 AA;

Query Match 100.0%; Score 258; DB 21; Length 60;

Best Local Similarity 100.0%; Pred. No. 1.3e-31;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGNVLINKDIRCKDDEFTHLTVLRPNTYEVKIDNSQVSGSLE 49

DB 12 NYKGNVLINKDIRCKDDEFTHLTVLRPNTYEVKIDNSQVSGSLE 60

RESULT 3

AAAY92352

ID AAAY92352 standard; Protein; 61 AA.

XX AC AAAY92352;

XX DT 10-AUG-2000 (first entry)

DE Recombinant human calreticulin residues 120-180.

XX MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;

KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;

KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;

KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.

OS Synthetic.

XX WO200020577-A1.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.

XX 06-OCT-1998; 98US-0103438.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Tosato G, Pike SE, Yao L;

XX WPI; 2000-303767/26.

XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,

XX useful for suppressing tumor growth

XX Claim 4; Page 82-83; 99pp; English.

XX A novel method of inhibiting endothelial cell growth comprises

XX contacting the cells with calreticulin (or its fragments/variants).

XX Fragments of calreticulin causes at least 40% inhibition of angiogenesis,

XX tumor growth and/or endothelial cell growth (claimed). The method may be

XX used for inhibiting angiogenesis in a patient. The angiogenesis is

XX associated with a disease other than a tumor that is associated with

XX neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,

XX trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune

XX inflammation, atherosclerosis, excessive wound repair, retinal

XX neovascularization, macular degeneration, corneal graft rejection,

XX contact lens overwear, Crohn's disease, non-immune inflammation,

CC rheumatoid arthritis, systemic lupus erythematosis, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 SQ Sequence 61 AA;

Query Match 100.0%; Score 258; DB 21; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.4e-31;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGNVLLINKDIRCKDETHLYTLVRPNTYEVKIDNSQVESGSLE 49
 |||||
 Db 13 nykgknvllinkdirckdethlytlvrpntyevkidsqvesgsle 61

RESULT 4
 AAY92351
 ID AAY92351 standard; Protein: 180 AA.
 XX
 AC AAY92351;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human vasostatin (calreticulin N-terminal 180 amino acids).
 XX

KW MBP-calreticulin; maltose binding protein; vasostatin; N-terminal;
 KW angiogenesis; inhibition; endothelial cell; anti-angiogenic;
 KW neuroprotective; antidiabetic; cytostatic; dermatological; hepatic;
 KW immunosuppressive; antiinflammatory; anti-atherosclerotic;
 KW gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.
 OS Synthetic.
 XX
 XX WO200020577-A1.
 XX
 XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.
 XX
 XX 06-OCT-1998; 98US-0103438.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Tosato G, Pike SE, Yao L;
 XX
 XX WPI; 2000-303767/26.

XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 XX
 XX Claim 4; Page 82; 99pp; English.

CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosis, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).

XX

SQ Sequence 180 AA;

Query Match 100.0%; Score 258; DB 21; Length 180;
 Best Local Similarity 100.0%; Pred. No. 5.7e-31;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGNVLLINKDIRCKDETHLYTLVRPNTYEVKIDNSQVESGSLE 49
 -|||||
 Db 132 nykgknvllinkdirckdethlytlvrpntyevkidsqvesgsle 180

RESULT 5
 AAY92355
 ID AAY92355 standard; Protein: 280 AA.
 XX
 AC AAY92355;

XX 10-AUG-2000 (first entry)
 XX
 DE Recombinant delta-120 calreticulin.
 XX

KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX Homo sapiens.
 OS Synthetic.
 XX
 XX WO200020577-A1.
 XX
 XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23240.
 XX
 XX 06-OCT-1998; 98US-0103438.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Tosato G, Pike SE, Yao L;
 XX
 XX WPI; 2000-303767/26.

XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 XX
 XX Claim 4; Page 86; 99pp; English.

XX This sequence comprises recombinant human calreticulin (AAY92350)
 CC missing the N-terminal 120 amino acids.
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The
 CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosis, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).

SQ Sequence 280 AA;

Query Match 100.0%; Score 258; DB 21; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1e-30;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGKVLINKDIRCKDDEFTHTLTVLRPDNTYEVKIDNSQVESGSL 49
 Db 12 NYKGKVLINKDIRCKDDEFTHTLTVLRPDNTYEVKIDNSQVESGSL 60

RESULT 6
 AAY92350
 ID AAY92350 standard; Protein; 400 AA.

XX AC AAY92350;
 XX DT 10-AUG-2000 (first entry)
 XX DE Recombinant human MBP-calreticulin.

XX KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatologic; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX OS Homo sapiens.
 XX PN WO200020577-A1.

XX PD 13-APR-2000.

XX PF 05-OCT-1999; 99WO-US23240.

XX PR 06-OCT-1998; 98US-0103438.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Tosato G, Pike SE, Yao L;

XX DR WPI; 2000-303767/26.

XX DR N-PSDB; AAA09346, AAA09347.

XX PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 XX useful for suppressing tumor growth

XX PS Claim 4; Page 80-81; 99pp; English.

XX CC Recombinant human MBP-calreticulin comprises the sequence of human
 CC calreticulin (see AAY92349) without the 17 N-terminal amino acids.
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The
 CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).

XX SQ Sequence 400 AA;

Query Match 100.0%; Score 258; DB 21; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1.6e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGKVLINKDIRCKDDEFTHTLTVLRPDNTYEVKIDNSQVESGSL 49
 Db 132 NYKGKVLINKDIRCKDDEFTHTLTVLRPDNTYEVKIDNSQVESGSL 180

RESULT 7
 AAW11156
 ID AAW11156 standard; peptide; 401 AA.

XX AC AAW11156;

XX DT 31-MAY-1997 (first entry)

XX DE Calreticulin.

XX KW calreticulin; C-domain; restenosis; inhibitor.

XX OS Homo sapiens.

XX PN WO9636643-A1.

XX PD 21-NOV-1996.

XX PF 17-MAY-1996; 96WO-IB00471.

XX PR 16-MAY-1996; 96US-0649417.

XX PR 17-MAY-1995; 95US-0442844.

XX PA (UYAL-) UNIV ALBERTA.

XX PI Lucas A, Michalak M;

XX DR WPI; 1997-012036/01.

XX PT Inhibition of restenosis in patients - using calreticulin or a
 XX C-domain polypeptide of calreticulin or a variant with the same
 XX activity.

XX PS Disclosure; Fig 1; 48pp; English.

XX CC The present sequence is calreticulin. It and a C-domain derived peptide
 CC (AAW06736) are useful for treating a patient to inhibit restenosis. The
 CC calreticulin-type cpds. are administered either parenterally,
 CC intravenously or via a catheter and can target areas of vascular damage
 CC to inhibit or prevent restenosis.

XX SQ Sequence 401 AA;

Query Match 100.0%; Score 258; DB 18; Length 401;
 Best Local Similarity 100.0%; Pred. No. 1.6e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGKVLINKDIRCKDDEFTHTLTVLRPDNTYEVKIDNSQVESGSL 49

Db 132 NYKGKVLINKDIRCKDDEFTHTLTVLRPDNTYEVKIDNSQVESGSL 180

RESULT 8

AAP92276

ID AAP92276 standard; protein; 417 AA.

XX AC AAP92276;

XX DT 23-FEB-1990 (first entry)

XX DE 60 kD Ro (Ro/SSA) antigen.

XX KW Sjorens syndrome; systemic lupus erythematosus.

XX OS Synthetic.

XX PN WO8909273-A.

XX PD 05-OCT-1989.

PF 22-MAR-1989; 89WO-US01213.
 XX
 PR 22-MAR-1988; 88US-0171634.
 XX
 PA (TEXA) UNIV OF TEXAS SYST.
 XX
 PI Sontheimer RD, Capra JD, McCauliffe DP;
 XX
 DR WPI; 1989-309537/42.
 DR N-PSDB; AAP92276.
 XX
 PT DNA sequences encoding antigenic epitope(s) of Ro 60 kD autoantigen
 PT - used in immunoassays to detect rheumatic disease
 XX
 PS Disclosure; Fig 2; 88pp; English.
 XX
 CC Synthetic peptides corresp. to an epitopic core of Ro antigen are
 CC expressed recombinantly to detect autoantibodies, for identification
 CC of autoimmune diseases. These epitopes are AAs 24-36, 23-36, 188-209,
 CC or 241-255. The peptides may be substd. for ribonucleoprotein particle
 CC antigens.
 XX
 SQ Sequence 417 AA;

Query Match 100.0%; Score 258; DB 10; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.7e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGKKNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 49
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 149 nygknvlinkdirckddefthlytlivrpontyevkidnsqvesgsle 197

RESULT 9
 AAY00927
 ID AAY00927 standard; Protein; 417 AA.
 XX
 AC AAY00927;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE Calreticulin.
 XX
 KW Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
 KW CUB functionality; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9907406-A1.
 XX
 PD 18-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-GB02430.
 XX
 PR 12-AUG-1997; 97GB-0016998.
 XX
 PA (UYLE-) UNIV LEICESTER.
 XX
 PI Schwaebler W;
 XX
 DR WPI; 1999-180404/15.
 XX
 PT Use of a cClqR binding domain - to modulate complement ubiquitin
 PT (CUB) functionality.
 XX
 PS Disclosure; Page 26-27; 31pp; English.
 XX
 CC This sequence is calreticulin, a homologue of Clq and collectin receptor
 CC (cClqR). The invention relates to the use of a cClqR binding domain in a

CC medicament to effect complement ubiquitin (CUB) functionality, and an
 CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 CC functionality. The cClqR binding domain, or its inhibitor, can be used to
 CC treat a human or animal body. Particularly an inhibitor is used to treat
 CC complement activation involved in the initiation and maintenance of
 CC inflammation, for example in myocardial infarction, brain ischaemia
 CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 CC erythematosus, burns, immune complex nephritis, and to treat amyloid
 CC plaques in Alzheimer's disease. The use of cClqR binding domain or
 CC inhibitor enables the CUB domain functionality to be modulated using a
 CC low molecular weight molecule.
 XX
 SQ Sequence 417 AA;

Query Match 100.0%; Score 258; DB 20; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.7e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGKKNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 49
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 149 nygknvlinkdirckddefthlytlivrpontyevkidnsqvesgsle 197

RESULT 10
 AAY92349
 ID AAY92349 standard; Protein; 417 AA.
 XX
 AC AAY92349;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human MBP-calreticulin.
 XX
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Peptide 1..17
 FT Protein /label= signal_peptide
 FT 18 /label= mature_protein
 XX
 PN WO200020577-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US23240.
 XX
 PR 06-OCT-1998; 98US-0103438.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Tosato G, Pike SE, Yao L;
 XX
 DR WPI; 2000-303767/26.
 DR N-PSDB; AAA09346, AAA09347.
 XX
 PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 XX
 PS Disclosure; Page 79-80; 99pp; English.
 XX
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis.
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with

CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angiofibromas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 SQ Sequence 417 AA;

Query Match 100.0%; Score 258; DB 21; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.7e-30;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 49
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 149 nykgknvlkndircckddefthlytlivrpdpntyevekidnsqvsge 197

RESULT 11
 AAW04171
 ID AAW04171 standard; Protein; 403 AA.

XX AAW04171;
 AC AAW04171;
 DT 12-DEC-1996 (first entry)
 XX
 DE Flea calreticulin PctCal403.
 XX
 KW Calreticulin; flea; haematophagous insect; allergic dermatitis;
 KW vaccine; therapy; PctCal403.
 XX
 OS Ctenocephalides felis.
 XX

PN W09628469-Al.

PD 19-SEP-1996.

PF 08-MAR-1996; 96WO-US03133.

PR 09-MAR-1995; 95US-0401509.

XX (HESK-) HESKA CORP.

XX Rushlow KE, Stiegler GL;

XX WPI; 1996-442861/44.

DR N-PSDB; AAT39516;

DR N-PSDB; AAT39517.

XX Haematophagous insect calreticulin protein - used to reduce insect

PT infestation and desensitize patients to allergic dermatitis

PS Claim 5; Page 68-69; 86pp; English.

XX Flea calreticulin protein PctCall589 (AAW04171) is a calcium-binding
 CC protein found in the salivary glands of Ctenocephalides felis.
 CC Its amino acid sequence was deduced from a cDNA clone (AAT39516)
 CC obtd. from a salivary gland cDNA library. Recombinant PctCall589
 CC can be produced in host cells transformed with a vector carrying
 CC calreticulin nucleic acids. Calreticulin alters the blood feeding
 CC behaviour of haematophagous insects and can be administered to an
 CC animal to reduce infestation. It reduces calreticulin activity in
 CC insects, so reducing the insect burden on an animal. Calreticulin
 CC can be used to elicit an immune response, thereby desensitising an
 CC animal to allergic dermatitis caused by fleas, mosquitoes or
 CC Culicoides.
 XX

SQ Sequence 403 AA;

Query Match 84.5%; Score 218; DB 17; Length 403;
 Best Local Similarity 79.6%; Pred. No. 1.8e-24;
 Matches 39; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 49
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 149 sykgknvlkndircckddefthlytlivrpdpntyevekidneksvgsnle 197

RESULT 12

AAW00924

ID AAY00924 standard; Protein; 122 AA.

XX AAY00924;

DT 28-MAY-1999 (first entry)

XX Human cClqR binding domain protein sequence.

XX Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
 KW CUB functionality; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy.
 XX

OS Homo sapiens.

XX WO9907406-Al.

PN 18-FEB-1999.

XX 12-AUG-1998; 98WO-GB02430.

XX 12-AUG-1997; 97GB-0016998.

XX (UYLE-) UNIV LEICESTER.

XX Schwaebler W;

DR WPI; 1999-180404/15.

DR N-PSDB; AAX27251.

XX Use of a cClqR binding domain - to modulate complement ubiquitin

PT (CUB) functionality.

XX Claim 9; Page 23; 31pp; English.

CC This sequence is a Clq and collectin receptor (cClqR) binding
 CC domain. The invention relates to the use of a cClqR binding domain in a
 CC medicament to effect complement ubiquitin (CUB) functionality, and an
 CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 CC functionality. The cClqR binding domain, or its inhibitor, can be used to
 CC treat a human or animal body. Particularly an inhibitor is used to treat
 CC complement activation involved in the initiation and maintenance of
 CC inflammation, for example in myocardial infarction, brain ischaemia
 CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 CC erythematosus, burns, immune complex nephritis, and to treat amyloid
 CC plaques in Alzheimer's disease. The use of cClqR binding domain or
 CC inhibitor enables the CUB domain functionality to be modulated using a
 CC low molecular weight molecule.
 XX

SQ Sequence 122 AA;

Query Match 73.6%; Score 190; DB 20; Length 122;
 Best Local Similarity 100.0%; Pred. No. 6.3e-21;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 14 RCKDDEFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 49
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 1 rckddefthlytlivrpdpntyevekidnsqvsge 36

RESULT 13

AA00926
 ID AAY00926 standard; Protein; 122 AA.
 XX AC
 XX AAY00926;
 DT 28-MAY-1999 (first entry)
 DE
 XX
 XX Rat cClqR binding domain protein sequence.
 XX
 XX C1q and collectin receptor; cClqR binding domain; complement ubiquitin;
 KW CUB functionality; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy.
 XX
 XX Rattus norvegicus.
 OS
 XX
 XX WO9907406-A1.
 PN
 XX
 XX 18-FEB-1999.
 PD
 XX
 XX 12-AUG-1998; 98WO-GB02430.
 PF
 XX
 XX 12-AUG-1997; 97GB-0016998.
 PR
 XX
 XX (UYLE-) UNIV LEICESTER.
 PA
 XX
 XX Schwaeble W;
 PI
 XX
 XX WPI: 1999-180404/15.
 XX
 XX N-PSDB; AAX27253.
 DR
 XX
 XX Use of a cClqR binding domain - to modulate complement ubiquitin
 PT (CUB) functionality.
 PT
 XX
 XX Claim 9; Page 24-25; 31pp; English.
 PS
 XX
 XX This sequence is a C1q and collectin receptor (cClqR) binding
 CC domain. The invention relates to the use of a cClqR binding domain in a
 CC medicament to effect complement ubiquitin (CUB) functionality, and an
 CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 CC functionality. The cClqR binding domain, or its inhibitor, can be used to
 CC treat a human or animal body. Particularly an inhibitor is used to treat
 CC complement activation involved in the initiation and maintenance of
 CC inflammation, for example in myocardial infarction, brain ischaemia
 CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 CC erythematosus, burns, immune complex nephritis, and to treat amyloid
 CC plaques in Alzheimer's disease. The use of cClqR binding domain or
 CC inhibitor enables the CUB domain functionality to be modulated using a
 CC low molecular weight molecule.
 XX
 XX Sequence 122 AA;
 SQ

Query Match 73.6%; Score 190; DB 20; Length 122;
 Best Local Similarity 100.0%; Pred. No. 6.3e-21;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 RCKDDEFTHLYTLVRPNTYEKIDNSQVESGSLE 49
 |||||||
 Db 1 rckddefthlytlvrpntyevkidsqvesgsle 36

RESULT 14

AA00925
 ID AAY00925 standard; Protein; 122 AA.
 XX AC
 XX AAY00925;
 XX
 DT 28-MAY-1999 (first entry)

XX
 DE
 XX
 KW
 KW
 KW
 KW
 KW
 OS
 OS
 PN
 XX
 XX
 PD
 XX
 PF
 XX
 XX
 PR
 XX
 XX
 PA
 XX
 XX
 PI
 XX
 DR
 DR
 XX
 XX
 PT
 PT
 XX
 PS
 XX
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 SQ

Mouse cClqR binding domain protein sequence.
 C1q and collectin receptor; cClqR binding domain; complement ubiquitin;
 CUB functionality; inhibitor; complement activation; inflammation;
 myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 immune complex nephritis; therapy.

Mus musculus.

WO9907406-A1.

18-FEB-1999.

12-AUG-1998; 98WO-GB02430.

12-AUG-1997; 97GB-0016998.

(UYLE-) UNIV LEICESTER.

Schwaeble W;

WPI: 1999-180404/15.

N-PSDB; AAX27252.

Use of a cClqR binding domain - to modulate complement ubiquitin
 (CUB) functionality.

Claim 9; Page 24; 31pp; English.

This sequence is a C1q and collectin receptor (cClqR) binding
 domain. The invention relates to the use of a cClqR binding domain in a
 medicament to effect complement ubiquitin (CUB) functionality, and an
 inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 functionality. The cClqR binding domain, or its inhibitor, can be used to
 treat a human or animal body. Particularly an inhibitor is used to treat
 complement activation involved in the initiation and maintenance of
 inflammation, for example in myocardial infarction, brain ischaemia
 (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 erythematosus, burns, immune complex nephritis, and to treat amyloid
 plaques in Alzheimer's disease. The use of cClqR binding domain or
 inhibitor enables the CUB domain functionality to be modulated using a
 low molecular weight molecule.

Sequence 122 AA;

Query Match 70.5%; Score 182; DB 20; Length 122;
 Best Local Similarity 97.2%; Pred. No. 1e-19;
 Matches 35; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 14 RCKDDEFTHLYTLVRPNTYEKIDNSQVESGSLE 49
 |||||||
 Db 1 rckddefthlytlvrpntyevkidsqvesgsle 36

RESULT 15

AA012312
 ID AAR12312 standard; Protein; 336 AA.
 XX AC
 XX AAR12312;
 XX
 DT 29-AUG-1991 (first entry)
 XX
 DE Partial sequence of Onchocera volvulus 42 kD antigen.
 XX
 KW River blindness; onchocerciasis; vaccine; antigen; parasite.
 XX
 OS Onchocerca volvulus.
 XX
 FH Key Location/Qualifiers
 FT Region 160..166

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2001, 03:00:05 ; Search time 76 Seconds
(without alignments)
319.074 Million cell updates/sec

Title: US-09-807-148-3
Perfect score: 2206
Sequence: 1 EPAYVFEQFLDGGTWRW.....EEDKEDEEDVPQAKDEL 400

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2206	100.0	400	21	Recombinant human
2	2206	100.0	417	10	60 kD Ro (Ro/SSA)
3	2206	100.0	417	21	Human MBP-calretic
4	2201	99.8	417	20	Calreticulin. Hom
5	2045.5	92.7	401	18	AAW11156
6	1533	70.4	280	21	Recombinant delta-
7	1450.5	65.8	403	17	Flea calreticulin
8	1235	56.0	336	12	Partial sequence o
9	1206.5	54.7	415	22	Castor bean calret
10	1206.5	54.7	415	22	AAAB66341
11	1164	52.8	421	21	Arabidopsis thalia

13	1164	52.8	424	21	AAAG24607	Arabidopsis thalia
13	1157	52.4	421	21	AAAG47932	Arabidopsis thalia
14	1157	52.4	424	21	AAAG47931	Arabidopsis thalia
15	1142.5	51.8	441	21	AAAG30997	Arabidopsis thalia
16	1142.5	51.8	444	21	AAAG30996	Arabidopsis thalia
17	1018.5	46.2	385	21	AAAB32385	Human secreted pro
18	971	44.0	180	21	AAAG2351	Human vasostatin (
19	907	41.1	312	21	AAAG24609	Arabidopsis thalia
20	900	40.8	312	21	AAAG47933	Arabidopsis thalia
21	885.5	40.1	332	21	AAAG30998	Arabidopsis thalia
22	799	36.2	593	16	AAAT1094	Calnexin sequence.
23	702	31.8	122	20	AAAY00924	Human cClqR bindin
24	699	31.7	122	20	AAAY00926	Rat cClqR binding
25	678	30.7	122	20	AAAY00925	Mouse cClqR bindin
26	658	29.8	530	21	AAAG26284	Arabidopsis thalia
27	658	29.8	530	21	AAAG46611	Arabidopsis thalia
28	641.5	29.1	542	22	AAAB66342	Castor bean calnex
30	606	27.5	532	21	AAAG04448	Arabidopsis thalia
31	606	27.5	548	21	AAAG04447	Arabidopsis thalia
32	532	24.1	394	21	AAAG26285	Arabidopsis thalia
33	532	24.1	394	21	AAAG46612	Arabidopsis thalia
34	510	23.1	394	21	AAAG04449	Arabidopsis thalia
35	510	23.1	417	21	AAAY77953	Arabidopsis thalia
36	446	20.2	85	17	AAW04170	A. thaliana enviro
37	351	15.9	70	21	AAAB44058	Flea calreticulin
38	339	15.4	274	21	AAAG26286	Human cancer assoc
39	327	14.8	61	21	AAAY92352	Arabidopsis thalia
40	318	14.4	60	21	AAAY92354	Recombinant human
41	297	13.5	119	21	AAAG44868	Recombinant human
42	258	11.7	49	21	AAAY92353	Zea mays protein f
43	250.5	11.4	84	21	AAAG41018	Recombinant human
44	230	11.3	515	21	AAAY77952	Zea mays protein f
45	246	11.2	235	21	AAAG50076	A. thaliana enviro
						Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAAY92350
ID AAAY92350 standard; Protein; 400 AA.
XX
AC AAAY92350;
XX
DT 10-AUG-2000 (first entry)
XX
XX Recombinant human MBP-calreticulin.
XX
XX R32-calreticulin; maltose binding protein; angiogenesis; inhibition;
KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
KW cytotstatic; dermalogical; immunosuppressive; antiinflammatory; hepatic;
KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
XX Homo sapiens.
XX
XX WO200020577-AI.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-US23240.
XX
XX 06-OCT-1998; 98US-0103438.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Iosato G, Pike SE, Yao L;
XX
XX WPI; 2000-303767/26.
XX
XX N-PSDB; AAA09346, AAA09347.
XX
XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
XX
XX useful for suppressing tumor growth

XX Claim 4; Page 80-81; 99pp; English.
 XX Recombinant human MBP-calreticulin comprises the sequence of human
 CC calreticulin (see AAY92349) without the 17 N-terminal amino acids.
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The
 CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX Sequence 400 AA:

Query Match 100.0%; Score 2206; DB 21; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1.8e-178;
 Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60
 Db 1 epavyfkeqfldgdgwtswieskhksdfgkflvssgkfygdeekdkgltsqdarfyal 60
 QY 61 SASFPFSSKNGQTLVVQFTVKHEQNDICGGYVKLFPPNSLDQTMHGDSEYNIMFGPDIC 120
 Db 61 sasfpfsskngqtlvvqftvkheqndicggvyvkflppnsldqtmhgdseynimfgpdic 120
 QY 121 GPGTKVHVIFNYKGNVLINKDIRCKDEFFHLYTLVIRPNTYEVKIDNSQVESGSLE 180
 Db 121 gpgtkvhvifnykgnvlinkdirckdeffhlytlvirpntyevkidnsqvesgsle 180
 QY 181 DWDDELPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDWDERAK 240
 Db 181 dwdelfppkkikdpdaskpedwderakidppdtskpedwderakidppdtskpedwderak 240
 QY 241 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPPDSIYAYDNFVGLGLD 300
 Db 241 dgewepvpiqnpeykgewkprqidnpdykgtwhpeidnpysppdsiyaydnfvglgld 300
 QY 301 LQOVKSGTIFDNFLITNDEAYAEFGNETGWVTKAAEKQMKQDEEORLKEEEDKKRK 360
 Db 301 lwovksgtifdnflitndeayaeefgnetgwvtkaeekqmkqdeeorlkeeedkkrrk 360
 QY 361 EEEAEADKEDDEDKDE 400
 Db 361 eeeaeaddkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedk 400

RESULT 2
 ID AAP92276 standard; protein; 417 AA.
 AC AAP92276;
 XX 23-FEB-1990 (first entry)
 DT 60 kD Ro (Ro/SSA) antigen.
 DE 60 kD Ro (Ro/SSA) antigen.
 XX Sjorens syndrome; systemic lupus erythematosus.
 KW Synthetic.
 OS Kaposi's sarcoma (claimed).
 XX W08909273-A.

XX 05-OCT-1989.
 XX 22-MAR-1989; 89WO-US01213.
 XX 22-MAR-1988; 88US-0171634.
 XX (TEXA) UNIV OF TEXAS SYST.
 XX Sontheimer RD, Capra JD, McCauliffe DP;
 XX WPI: 1989-309537/42.
 XX N-PSDB; AAP92276.
 XX DNA sequences encoding antigenic epitope(s) of Ro 60 kD autoantigen
 PT - used in immunoassays to detect rheumatic disease
 XX Disclosure: Fig 2; 88pp; English.
 XX Synthetic peptides corresp. to an epitopic core of Ro antigen are
 CC expressed recombinantly to detect autoantibodies, for identification
 CC of autoimmune diseases. These epitopes are AAs 24-36, 23-36, 188-209,
 CC or 241-255. The peptides may be substd. for ribonucleoprotein particle
 CC antigens.
 XX Sequence 417 AA:

Query Match 100.0%; Score 2206; DB 10; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.9e-178;
 Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60
 Db 18 epavyfkeqfldgdgwtswieskhksdfgkflvssgkfygdeekdkgltsqdarfyal 77
 QY 61 SASFPFSSKNGQTLVVQFTVKHEQNDICGGYVKLFPPNSLDQTMHGDSEYNIMFGPDIC 120
 Db 78 sasfpfsskngqtlvvqftvkheqndicggvyvkflppnsldqtmhgdseynimfgpdic 137
 QY 121 GPGTKVHVIFNYKGNVLINKDIRCKDEFFHLYTLVIRPNTYEVKIDNSQVESGSLE 180
 Db 138 gpgtkvhvifnykgnvlinkdirckdeffhlytlvirpntyevkidnsqvesgsle 197
 QY 181 DWDDELPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDWDERAK 240
 Db 198 dwdelfppkkikdpdaskpedwderakidppdtskpedwderakidppdtskpedwderak 257
 QY 241 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPPDSIYAYDNFVGLGLD 300
 Db 258 dgewepvpiqnpeykgewkprqidnpdykgtwhpeidnpysppdsiyaydnfvglgld 317
 QY 301 LQOVKSGTIFDNFLITNDEAYAEFGNETGWVTKAAEKQMKQDEEORLKEEEDKKRK 360
 Db 318 lwovksgtifdnflitndeayaeefgnetgwvtkaeekqmkqdeeorlkeeedkkrrk 377
 QY 361 EEEAEADKEDDEDKDE 400
 Db 378 eeeaeaddkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedk 417

RESULT 3
 ID AAP92349 standard; protein; 417 AA.
 AC AAP92349;
 XX 10-AUG-2000 (first entry)
 DT Human MBP-calreticulin.
 DE MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 XX endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW

KW cytotatic; dermatological; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Peptide 1..17
 FT /label= signal_peptide
 FT 18
 FT /label= mature_protein
 XX
 PN WO200020577-A1.
 XX
 XX
 PD 13-APR-2000.
 XX
 XX 05-OCT-1999; 99WO-US23240.
 XX
 XX 06-OCT-1998; 98US-0103438.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Tosato G, Pike SE, Yao L;
 XX WPI: 2000-303767/26.
 DR N-PSDB; AAA09346, AAA09347.
 DR
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 PT
 XX Disclosure; Page 79-80; 99pp; English.

XX A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of angiogenesis,
 CC tumor growth and/or endothelial cell growth (claimed). The method may be
 CC used for inhibiting angiogenesis in a patient. The angiogenesis is
 CC associated with a disease other than a tumor that is associated with
 CC neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia,
 CC trachoma, neovascular glaucoma, psoriasis, angioblastomas, immune
 CC inflammation, atherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythematosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 XX Sequence 417 AA;

Query Match 100.0%; Score 2206; DB 21; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.9e-178;
 Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPAYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKGLQTSQDARYAL 60
 DB 18 epayfkeqfldgdgwtswieskhksdfgkflvssgkfygdeekdglqtsqdarfyal 77
 QY 61 SASFEPFSNKGQTLVQFTVKHEQNDICGGYVKLPNSLDQTDHMGDSYNTMFGPDIC 120
 DB 78 sasfepfsnkgqtlvqftvkheqndicgggyvklpnsldqtdmngdsyntmfgpdic 137
 QY 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVESGSL 180
 DB 138 gpcstkvhvifnykgnvlinkdirckddethlytlivrpdntyevkidnsqvesgsl 197
 QY 181 DDWDFLPPKKIKDPDASKPDPEDWERAIDDPDTSKPEDWPKPHIIPDPDAKKPDDWDEEM 240
 DB 198 ddwdfllppkkikdpdaskpdpedweraidpdpdtskpedwkpheipdpdakkpddwdeem 257
 QY 241 DGEWEPPVIONPYKGEWKPRQIDNPDYKGTWTHPEIDNPEYSPDPSIAYDNFGVLGLD 300
 DB 18 dgeweppvionpykgewkprqidnpykgtwthpeidnpeyspdpsiyaydnfgvlgld 300

Db 258 dgeweppvionpykgewkprqidnpykgtwthpeidnpeyspdpsiyaydnfgvlgld 317
 QY 301 LWQVKSGTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKQDEQRKLEEEEDKKRK 360
 DB 318 lwqvksgtifdnflitndeayaeefgnetgwtkaaekqmkdkqdeeqrikeeedkkrk 377
 QY 361 EEEAEADKDEDEKDE 400
 DB 378 eeeaeadkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedkdedk 417

RESULT 4

AA00927

ID AAY00927 standard; Protein; 417 AA.

XX AC AAY00927;

XX DT 28-MAY-1999 (first entry)

XX DE Calreticulin.

XX Clq and collectin receptor; cClqR binding domain; complement ubiquitin;
 KW CUB functionality; inhibitor; complement activation; inflammation;
 KW myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;
 KW rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;
 KW immune complex nephritis; therapy.

XX OS Homo sapiens.

XX PN WO9907406-A1.

XX PD 18-FEB-1999.

XX PF 12-AUG-1998; 98WO-GB02430.

XX PR 12-AUG-1997; 97GB-0016998.

XX PA (UYLE-) UNIV LEICESTER.

XX PI Schwaebler W;

XX DR WPI: 1999-180404/15.

XX PT Use of a cClqR binding domain - to modulate complement ubiquitin
 XX {CUB} functionality.
 XX PS Disclosure; Page 26-27; 31pp; English.

XX CC This sequence is calreticulin, a homologue of Clq and collectin receptor

XX {cClqR}. The invention relates to the use of a cClqR binding domain in a
 CC medicament to effect complement ubiquitin (CUB) functionality, and an
 CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB
 CC functionality. The cClqR binding domain, or its inhibitor, can be used to
 CC treat a human or animal body. Particularly an inhibitor is used to treat
 CC complement activation involved in the initiation and maintenance of
 CC inflammation, for example in myocardial infarction, brain ischaemia
 CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus
 CC erythematosus, burns, immune complex nephritis, and to treat amyloid
 CC plaques in Alzheimer's disease. The use of cClqR binding domain or
 CC inhibitor enables the CUB domain functionality to be modulated using a
 CC low molecular weight molecule.

XX SQ Sequence 417 AA;

Query Match 99.8%; Score 2201; DB 20; Length 417;

Best Local Similarity 99.8%; Pred. No. 5.1e-178;

Matches 399; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPAYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKGLQTSQDARYAL 60

DB 18 epayfkeqfldgdgwtswieskhksdfgkflvssgkfygdeekdglqtsqdarfyal 77

QY 61 SASFEPSNKGOTLVQFTVKHEONIDCGGYVKKLPNSLDQTDHMGDSEYNIMFGPDIC 120
 DB 78 sasfepsnkgqclvqftvkheqndcgggyvkklpnsldqtdmhgdseynimfgpdic 137
 QY 121 GPQTKVHVIFNYKGNVLINIDIRCKODEFTHLYTLIVRPDNTYEVKIDNSOVESGSLE 180
 DB 138 gpqtkvhvifnykgnvlinkdirckddefthlytlivrpndntevkldnsqvesgsle 197
 QY 181 DWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDASKPEDWDERAK 240
 DB 198 ddwdfippkkikdpdaskpedwderakidpdpdtskpedwderakidpdpdaskpedwderak 257
 QY 241 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPHIDNPESPDPSIYAYDNFGVLGLD 300
 DB 258 dgewepvtonpeykgewkprqidnpdkygtwhphidnpespdpsiyaydnfvgvlgid 317
 QY 301 LMQVKSCTIFDNFLITNDEAYAEFGNETGWVTKAAEKQMKQKQDEQRLKEEEDKKR 360
 DB 318 lmqvksctifdnflitndeayaeefnetgwvtkaekqmkdkqdeeqrlkeeedkkrrk 377
 QY 361 EEEEAEDKEDDE 400
 DB 378 eeeeaedkdedkdedeedeedeedeedeedeedeedeedeedeedeedeedeedeede 417

RESULT 5

AAW11156
 ID AAW11156 standard; peptide; 401 AA.
 AC AAW11156;
 XX
 DT 31-MAY-1997 (first entry)
 XX
 DE Calreticulin.
 XX
 KW calreticulin; C-domain; restenosis; inhibitor.
 XX
 OS Homo sapiens.
 XX
 PN W09636643-Al.
 XX
 PD 21-NOV-1996.
 XX
 PF 17-MAY-1996; 96WO-IB00471.
 XX
 PR 16-MAY-1996; 96US-0649417.
 PR 17-MAY-1995; 95US-0442844.
 XX
 PA (UYAL-) UNIV ALBERTA.
 XX
 XX Lucas A, Michalak M;
 PI WPI; 1997-012036/01.
 DR
 XX
 PT Inhibition of restenosis in patients - using calreticulin or a
 PT C-domain polypeptide of calreticulin or a variant with the same
 PT activity.
 XX
 PS Disclosure; Fig 1; 48pp; English.
 XX
 CC The present sequence is calreticulin. It and a C-domain derived peptide
 CC (AAW06736) are useful for treating a patient to inhibit restenosis. The
 CC calreticulin-type cpds. are administered either parenterally,
 CC intravenously or via a catheter and can target areas of vascular damage
 CC to inhibit or prevent restenosis.
 XX
 SQ Sequence 401 AA;

Query Match 92.7%; Score 2045.5; DB 18; Length 401;
 Best Local Similarity 92.5%; Pred. No. 6.7e-165;
 Matches 371; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

QY 1 EPAVYFKQFLDGDGTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60
 DB 1 epvvyfkeqfldgdgtsrwieskhksdfgkflvssgkfygddekdkglqtsqdarfyal 60
 QY 61 SASFEPSNKGOTLVQFTVKHEONIDCGGYVKKLPNSLDQTDHMGDSEYNIMFGPDIC 120
 DB 61 sarfepsnkgqclvqftvkheqndcgggyvkklpnsldqtdmhgdseynimfgpdic 130
 QY 121 GPQTKVHVIFNYKGNVLINIDIRCKODEFTHLYTLIVRPDNTYEVKIDNSOVESGSLE 180
 DB 121 gpqtkvhvifnykgnvlinkdirckddefthlytlivrpndntevkldnsqvesgsle 190
 QY 181 DWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDASKPEDWDERAK 240
 DB 181 ddwdfippkkikdpdaskpedwderakidpdpdtskpedwderakidpdpdaskpedwderak 240
 QY 241 DGEWEPVTONPEYKGEWKPRQIDNPDKYKGTWHPHIDNPESPDPSIYAYDNFGVLGLD 300
 DB 241 dgewepvtonpeykgewkprqidnpdkygtwhphidnpespdpsiyaydnfvgvlgid 300
 QY 301 LMQVKSCTIFDNFLITNDEAYAEFGNETGWVTKAAEKQMKQKQDEQRLKEEEDKKR 360
 DB 301 lmqvksctifdnflitndeayaeefnetgwvtkaekqmkdkqdeeqrlkeeedkkrrk 360
 QY 361 EEEEA-EDKEDDE 400
 DB 361 eeeeaedkdedkdedeedeedeedeedeedeedeedeedeedeedeedeedeedeede 401

RESULT 6

AAW11156
 ID AAW11156 standard; Protein; 280 AA.
 AC AAW11156;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Recombinant delta-120 calreticulin.
 XX
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;
 KW cytostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN W0200020577-Al.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US23240.
 XX
 PR 06-OCT-1998; 98US-0103438.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Tosato G, Pike SE, Yao L;
 DR WPI; 2000-303767/26.
 XX
 PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,
 PT useful for suppressing tumor growth
 XX
 PS Claim 4; Page 86; 99pp; English.
 XX
 CC This sequence comprises recombinant human calreticulin (AAW92350)
 CC missing the N-terminal 120 amino acids.
 CC A novel method of inhibiting endothelial cell growth comprises
 CC contacting the cells with calreticulin (or its fragments/variants).
 CC Fragments of calreticulin causes at least 40% inhibition of
 CC angiogenesis, tumor growth and/or endothelial cell growth (claimed). The
 CC method may be used for inhibiting angiogenesis in a patient. The

CC angiogenesis is associated with a disease other than a tumor that is
 CC associated with neovascularization (e.g. diabetic neuropathy, retrolental
 CC fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiofibromas,
 CC immune inflammation, attherosclerosis, excessive wound repair, retinal
 CC neovascularization, macular degeneration, corneal graft rejection,
 CC contact lens overwear, Crohn's disease, non-immune inflammation,
 CC rheumatoid arthritis, systemic lupus erythromatosus, thyroiditis,
 CC Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's
 CC syndrome, sarcoidosis and primary biliary cirrhosis). The method may
 CC also be used for treating/inhibiting tumor growth especially
 CC Kaposi's sarcoma (claimed).
 XX
 XX Sequence 280 AA;

Query Match 70.4%; Score 1553; DB 21; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1.8e-123;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GPCPKVHVIFNFKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 180
 DB 1 gpgckkvhvifnfykgkgnvinkdirckddefthlytlivrpdtntyevkdnsgvesgsl 60
 QY 181 DDMDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDPKPEHIPDPAKKPDWDEEM 240
 DB 61 ddwdfllppkikdpdaskpedwderakiddptdskpedwkppehipdpdakkpedwdeem 120
 QY 241 DGEWEPVIONPEYKGEWKPRQIDNPYKGTWTHPEIDNPEYSPDPSIYAYDNFVGLGLD 300
 DB 121 dgewepvionpeykgewkprqidnpykgtwthpeidnpeyspdpsiyaydnfvgvlgld 180
 QY 301 LWQVKSCTIFDNFLITNDEAYAEFGNETWGVTKAAEQMKDKQDEORLKEEEDKKR 360
 DB 181 lwqvksctifdnflitndeayaeefgnwtgvtkaaeqmkdkqdeeqrlkeeedkkrr 240
 QY 361 EEEAEKDE 400
 DB 241 eeeaeckdedekdedekdedekdedekdedekdedekdedekdedekdedekdedekded 280

RESULT 7
 ID AAW04171
 XX AAW04171 standard; Protein; 403 AA.
 AC AAW04171;
 XX
 XX 12-DEC-1996 (first entry)
 DT
 XX Flea calreticulin Pctcal403.
 DE
 XX Calreticulin; flea: haematophagous insect; allergic dermatitis;
 KW vaccine; therapy; Pctcal403.
 XX
 XX Ctenocephalides felis.
 OS
 XX WO9628469-A1.
 PN
 XX 19-SEP-1996.
 PD
 XX
 XX 08-MAR-1996; 96WO-US03133.
 XX
 XX 09-MAR-1995; 95US-0401509.
 XX
 XX (HESK-) HESKA CORP.
 PA
 XX Rushlow KE, Stiegler GL;
 XX
 XX WPI; 1996-442861/44.
 DR N-PSDB; AAT39516;
 DR N-PSDB; AAT39517.
 XX
 XX Haematophagous insect calreticulin protein - used to reduce insect
 PT infestation and desensitise patients to allergic dermatitis

XX
 XX Ciaim 5; Page 68-69; 86pp; English.
 XX
 CC Flea calreticulin protein Pctcal1589 (AAW04171) is a calcium-binding
 CC protein found in the salivary glands of Ctenocephalides felis.
 CC 1-5 amino acid sequence was deduced from a cDNA clone (AAT39516)
 CC obtd from a salivary gland cDNA library. Recombinant Pctcal1589
 CC can be produced in host cells transformed with a vector carrying
 CC calreticulin nucleic acids. Calreticulin alters the blood feeding
 CC behaviour of haematophagous insects and can be administered to an
 CC animal to reduce infestation. It reduces calreticulin activity in
 CC insects, so reducing the insect burden on an animal. Calreticulin
 CC can be used to elicit an immune response, thereby desensitising an
 CC animal to allergic dermatitis caused by fleas, mosquitoes or
 CC Culicoides.
 XX
 XX Sequence 403 AA;

Query Match 65.8%; Score 1450.5; DB 17; Length 403;
 Best Local Similarity 67.8%; Pred. No. 1.4e-114;
 Matches 265; Conservative 44; Mismatches 71; Indels 11; Gaps 6;

QY 2 PAVYKEQFIDGQWTSRWIESKHK-SDFGKFVLSGKFGYDEKDKGLOTSDARFVAL 60
 DB 19 pevleefnfvd-dtwtntwysehpgefkgfvtgtagkfyndaeadkglgtgqdarfyal 77
 QY 61 SASPEFSNKGQTLVQFTYKHEQNICDGGYVKKLFNSLDQDTMHGDSYINIMFGPDIC 120
 DB 78 shkfkpsnkdktlivqfsvkhegnidcgggykkgfsvnqkdmhgespyelmfgpdic 137
 QY 121 GPCTKKVHVIFNFKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 180
 DB 138 dpgtkkvhvifnfykgkgnvinkdirckddvthvtytlivkpdntyevlidnekvesgnle 197
 QY 181 DDMDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDPKPEHIPDPAKKPDWDEEM 240
 DB 198 ddweflapkikdpdaskpedwderatidpndtkpedwkppehipdpdakkpedwdeem 257
 QY 241 DGEWEPVIONPEYKGEWKPRQIDNPYKGTWTHPEIDNPEYSPDPSIYAYDNFVGLGLD 300
 DB 358 dgewepamidnpeykgewkprqidnpykgtwvhwpeidnpeypdtqlkrdcaigld 317
 QY 301 LWQVKSCTIFDNFLITNDEAYAEFGNETWGV--TKAAEQMKDKQDEORLKEEEDKK 358
 DB 318 lwqvksctifdnflitndeayaeefgnwtgvtkaaeqmkdkqdeeqrlkeeedkkaree--- 371
 QY 359 RKEEEAEKDED-DE 388
 DB 372 tkeendtdadedldnaetpedtldaehde 402

RESULT 8
 ID AAR12312
 XX AAR12312 standard; Protein; 336 AA.
 AC AAR12312;
 XX
 XX 29-AUG-1991 (first entry)
 DT
 XX Partial sequence of Onchocera volvulus 42 kD antigen.
 DE
 XX River blindness; onchocerciasis; vaccine; antigen; parasite.
 KW Onchocerca volvulus.
 XX
 XX Location/Qualifiers
 FT Region 160..166
 FT /label= repeat unit
 FT /note= "hydrophilic"
 FT Region 177..183
 FT /label= repeat unit
 FT /note= "hydrophilic"

FT Region 195..201
 FT /label= repeat unit
 FT /note= "hydrophilic"
 PN US021342-A.
 PD 04-JUN-1991.
 XX 30-JUN-1988; 88US-0214264.
 PF 30-JUN-1988; 88US-0214264.
 PR (UYHO-) UNIV HOSPITALS CLEV.
 XX
 PI Greene BM, Unnasch TR;
 XX
 DR WPI; 1991-185179/25.
 DR N-PSDB; AAQ11987.
 XX
 PT DNA encoding Onchocerca volvulus antigen - used to express
 PT recombinant antigen for vaccine against onchocerciasis or river
 PT blindness.
 XX
 PS Disclosure; Fig 7; 20pp; English.
 XX
 CC The sequence was deduced from a cDNA clone lambda RAL-1 prep.
 CC from RNA isolated from modules excised from patients infected
 CC with O. volvulus. The N-terminal is incomplete, however Abs
 CC which specifically bind to protein prep. from induced cultures
 CC of lambda RAL-1 lysogens recognise a single polypeptide of mol.
 CC wt. 42,000 in extracts of adult worms. Analysis deduced sequence
 CC suggests that it encodes a protein of mol. wt. 39,130. If the
 CC antigen is not subject to post-translational processing this
 CC suggests that most of the coding sequence is present. The three
 CC repeats are highly hydrophilic regions likely to be exposed on
 CC the surface of the antigen and highly immunogenic. Recombinant
 CC antigen expressed by the clone can be used stimulate T-cells of
 CC individuals infected by the parasite to proliferate and may be
 CC used as the basis for a vaccine against Onchocerciasis or river
 CC blindness.
 XX
 SQ Sequence 336 AA;

Query Match 56.0%; Score 1235; DB 12; Length 336;
 Best Local Similarity 64.7%; Pred. No. 1.8e-96;
 Matches 218; Conservative 48; Mismatches 63; Indels 8; Gaps 3;
 QY 39 FYGDEEKGLQTSODAREYALSASFE-PFSNKGQTLVQVFTVKHEQNIDCGGYYKLP 97
 Db 1 fygdavkdkgkttgdaqfysigaktdksfnskgkslvigfsvkheqdcgggyvklma 60
 QY 98 NSLDQDMHGDSEYNIIMFGPDICGPGTKKVVHVFYNYKGNVLNKIDIRCKDDEFTHLYTL 157
 Db 61 sdvnledsgetpyhmfpgdicgpgtkkvfvfhykdrhmkkairckdadvfthlytl 120
 QY 158 IVRPDNTYEVKIDNSOVESGSLDDWDFLPPKKIKDPDASKPEDMDERAKIDDPDTSKPE 217
 Db 121 ivnsdnteyvqidgaesgeleadwdfipppkkikdpdaskpdedwerefiddeddkpe 180
 QY 218 DWDKPEHIPPDAKKPEDWDEMGEPVTONPEYKGEWKPQIDNPDKYKGTWHPHPI 277
 Db 181 dwdkpehirdpdkkpedwddmgewepmvdnpeykgewkpkqknpaykxgwhpei 240
 QY 278 DNPESPDPSIAYDNFGLDLMOVKSGTIFDNLITNDAYAEFGNETGWVTKAAE 337
 Db 241 eipdytpdanlyyddgaigldwqykgstifddvdividsveeakkfgeklktitreg 300
 QY 338 KONKQKDEEQLKEEED-----KKRKEEAEADK 368
 Db 301 kk-kgkktkqkkeknekikkekmmkrkrankkkk 336

RESULT 9
 AAB6341
 ID AAB66341 standard; Protein; 415 AA.
 XX AC
 XX AAB66341;
 XX
 DT 05-APR-2001 (first entry)
 XX
 DE Castor bean calreticulin SEQ ID NO: 2.
 XX
 KW Castor bean; calreticulin; calcium binding; er; endoplasmic reticulum;
 KW calnexin; promoter; resistance.
 XX
 OS Ricinus communis.
 XX
 PN US6171864-B1.
 XX
 PD 09-JAN-2001.
 XX
 PF 05-JUL-1996; 96US-0675816.
 XX
 PR 05-JUL-1996; 96US-0675816.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Coughlan SJ, Winfrey RJ;
 XX
 DR WPI; 2001-122335/13.
 DR N-PSDB; AAF29741, AAF29742.
 XX
 PT New nucleic acid molecules encoding a calcium binding chaperone protein
 PT in endoplasmic reticulum, calreticulin, and calreticulin promoter
 PT sequences, useful for producing foreign gene products in plant cells -
 XX
 PS Claim 1; Fig 3; 45pp; English.
 XX
 CC The present invention provides the protein and coding sequences for the
 CC castor bean calreticulin protein and the calreticulin promoter sequence.
 CC In addition, the castor bean calnexin protein, coding sequence and
 CC promoter are also described. Calreticulin and calnexin are calcium
 CC binding proteins found in the endoplasmic reticulum (er). The
 CC calreticulin promoter sequence can be used in vectors to promote the
 CC expression of foreign genes, particularly resistance genes, in plant
 CC cells.
 XX
 SQ Sequence 415 AA;

Query Match 54.7%; Score 1206.5; DB 22; Length 415;
 Best Local Similarity 55.7%; Pred. No. 6.1e-94;
 Matches 225; Conservative 56; Mismatches 106; Indels 17; Gaps 7;
 QY 4 VYFKEQFLDGDGWTSRWIESKHKSQ---PGKFVLSGKFGYGDDEKDKGLQTSQDARFYAL 60
 Db 22 vifeerf--edgwenrwwksdwdkddentagewnytskgwngd-pndkgigtseydyfyal 78
 QY 61 SASFEPFSGKTLVQVFTVKHEQNIDCGGYYKLPNSLDQDMHGDSEYNIIMFGPDIC 120
 Db 79 saefpfsnkdktlvfqsvkheqkldcggymkllssstdgkkgfgdtpyimfgpdic 138
 QY 121 GGTGKVVHVFYNYKGNVLNKIDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSOVESGSL 180
 Db 139 gystkkvhailnyndtnhlikkevpcetdqlthvylvirpdatsilidnvektqgsly 198
 QY 181 DWDKPEHIPPDAKKPEDWDEMGEPVTONPEYKGEWKPQIDNPDKYKGTWHPHPI 239
 Db 199 tdwdlppkkikdpdkkpedwdekeyipdpdkkpegdyddipkelpdpdkkpedwde 258
 QY 240 MDGEWEPVTONPEYKGEWKPQIDNPDKYKGTWHPHPIIDNPESPDPSIAYDNFVGL 299
 Db 259 edgeweptianpeykgwppkpkiknpnykgwkapmidnfpdkdpelvyvynlkyvqi 318
 QY 300 DLWQVKSQTIFDNLITNDAYAEFGNETGWVTKAAEQKMDKQDEORLKEEEDKKR 359

Db	319	elwvqsgtllfndvncldonpeyaklaetwgnkdaeka---afeaeakkeeeekdd	375
Qy	360	KEEEAEADKEDDEKDEDEDEDEKDE---EDEEDVPQAKDEL	400
Db	376	padsdadedddaddtegedgdsdaedaedv---hdel	415
RESULT	10		
ID	AAB66343		
AC	AAB66343		
DT	05-APR-2001 (first entry)		
DE	Castor bean calreticulin.		
KW	Castor bean; calreticulin; calcium binding; er; endoplasmic reticulum;		
KW	calnexin; promoter; resistance.		
OS	Ricinus communis.		
PN	US6171864-B1.		
PD	09-JAN-2001.		
XX	05-JUL-1996; 96US-0675816.		
XX	05-JUL-1996; 96US-0675816.		
XX	(PION-) PIONEER HI-BRED INT INC.		
XX	Coughlan SJ, Winfrey RJ;		
XX	WPI; 2001-122335/13.		
XX	N-PSDB; AAF29755.		
PT	New nucleic acid molecules encoding a calcium binding chaperone protein		
PT	in endoplasmic reticulum, calreticulin, and calreticulin promoter		
PT	sequences, useful for producing foreign gene products in plant cells		
PS	Disclosure; Fig 1B; 45pp; English.		
XX			
CC	The present invention provides the protein and coding sequences for the		
CC	castor bean calreticulin protein and the calreticulin promoter sequence		
CC	In addition, the castor bean calnexin protein, coding sequence and		
CC	promoter are also described. Calreticulin and calnexin are calcium		
CC	binding proteins found in the endoplasmic reticulum (er). The		
CC	calreticulin promoter sequence can be used in vectors to promote the		
CC	expression of foreign genes, particularly resistance genes, in plant		
CC	cells.		
XX			
SQ	Sequence 415 AA;		
Query Match	54.7%;	Score 1206.5;	DB 22; Length 415;
Best Local Similarity	55.7%;	Pred. No. 6,1e-94;	
Matches	225;	Conservative 56;	Mismatches 106; Indels 17; Gaps
Qy	4	VYPKEQFDGDMGWTNRWIESKHSB---FGKPVLSGGKFGYGDDEKGLQTSQDAFYAL	60
Db	22	vffeef--edgwenrwnksdwkkddentagewnyLsgkwngd-pndkqigtseayrfyal	78
Qy	61	SASFEPFSNKGQTLVQVPTVKHEQNIIDCGGVYKLFPPNSLQDTMHDGSEYNVFGPDIC	120
Db	79	saeefpfsnkdktlvffgsvkheqldcgggymkllssstdqkkgfggtpysinfgpdic	138
Qy	121	GPCTKKVHVFNKGNVNLINKDCKDEFTHTLYTLVRPNTYEVKIDNSQVESGSLE	180
Db	139	gystkvvhalnyndcnhlikvepctdqlChvylvirpdatysllidnvektgtsly	198
Qy	181	DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSDXPEDWDK-PHPHIPPDDAKKPEDWDEE	239

Db	199	tdwdllppkikdpkaakpewdekeyipdpedkkgpyddipkeipdpdaakpewddde	258
Qy	240	MDGEWEPPIQNPEYKGEWKPRQIDNPDPYKGTWIHPIDNPESPDPSIYAYDNFVGLGL	299
Db	259	edgewaptianpeykpwpkpkikpnpykgwkapmidnpdkfddpelyvynlkyyvgi	318
Qy	300	DLHQVKGSTPDNPLIINDRAYAEFGNETWGVTKAAEKOMKQDQFEORLKEEEDKKR	359
Db	319	elwqvgstgldfnvllcndpeyakqlaetwgnkndaeka---afeaeekkkkeeeskdd	375
Qy	360	KEEEAEEDKDEDEDEDEDEDEKE---EDEEDVPGQAKDEL	400
Db	376	padsdadaddadddtegedgesksdaaedsaedv---hdel	415
RESULT	11		
AA	AAG24608		
ID	AAG24608	standard; Protein; 421 AA.	
XX	AC		
XX	AC	AAG24608;	
DT	17-OCT-2000	(first entry)	
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 28350.	
KW	KW	Protein identification; signal transduction pathway; metabolic pathwa	
KW	KW	hybridisation assay; genetic mapping; gene expression control; promot	
KW	KW	termination sequence.	
XX	OS	Arabidopsis thaliana.	
XX	PN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF	25-FEB-2000; 2000EP-0301439.	
XX	PR	25-FEB-1999; 99US-0121825.	
PR	PR	05-MAR-1999; 99US-0123180.	
PR	PR	09-MAR-1999; 99US-0123548.	
PR	PR	23-MAR-1999; 99US-0125788.	
PR	PR	25-MAR-1999; 99US-0126264.	
PR	PR	29-MAR-1999; 99US-0126785.	
PR	PR	01-APR-1999; 99US-0127462.	
PR	PR	06-APR-1999; 99US-0128234.	
PR	PR	08-APR-1999; 99US-0128714.	
PR	PR	16-APR-1999; 99US-0129845.	
PR	PR	19-APR-1999; 99US-0130077.	
PR	PR	21-APR-1999; 99US-0130449.	
PR	PR	23-APR-1999; 99US-0130510.	
PR	PR	23-APR-1999; 99US-0130891.	
PR	PR	28-APR-1999; 99US-0131449.	
PR	PR	30-APR-1999; 99US-0132048.	
PR	PR	30-APR-1999; 99US-0132407.	
PR	PR	04-MAY-1999; 99US-0132484.	
PR	PR	05-MAY-1999; 99US-0132485.	
PR	PR	06-MAY-1999; 99US-0132486.	
PR	PR	06-MAY-1999; 99US-0132487.	
PR	PR	07-MAY-1999; 99US-0132863.	
PR	PR	11-MAY-1999; 99US-0134256.	
PR	PR	14-MAY-1999; 99US-0134218.	
PR	PR	14-MAY-1999; 99US-0134219.	
PR	PR	14-MAY-1999; 99US-0134221.	
PR	PR	14-MAY-1999; 99US-0134370.	
PR	PR	18-MAY-1999; 99US-0134768.	
PR	PR	19-MAY-1999; 99US-0134941.	
PR	PR	20-MAY-1999; 99US-0135124.	
PR	PR	21-MAY-1999; 99US-0135353.	
PR	PR	24-MAY-1999; 99US-0135629.	
PR	PR	25-MAY-1999; 99US-0136021.	
PR	PR	27-MAY-1999; 99US-0136392.	
PR	PR	28-MAY-1999; 99US-0136782.	

3

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139465.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match          52.8% Score 1164; DB 21; Length 421;
Best Local Similarity 54.2%; Pred. No. 2.4e-90;
Matches 221; Conservative 57; Mismatches 114; Indels 16; Gaps 8;

QY 3 AVYKEQFLDGDWTSRWIESKHKSD---FGKFLVSSGKFGYDEEKDKGLQTSQDARFYA 59
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 20 avifeerf--ddgwenrwwksekddntagewkhtagnwsgd-andkgqtseydyfya 76
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 60 LSASFEPFSNKGQTLVVQFVTKHEQNTDCGGYVVKLFPSNLDQDTHMGDSEYNTMFGPDI 119
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 77 isaeifelsnkdktlvfqsavkheqldcggygmklsgdvdkkfggdtphyslmfgpdi 136
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 120 CGPGTKVHVIFNFKGNVLINKDIRCKDDEFTHLTVLRPONTYEVKINDSOVESGSL 179
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 137 cgystkvhailtneanhlkkdvpctdqtlthvffilrpdatsyllndnvektqgs 196
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 180 EDDWDFLPPKKIKDPDASKEDWDERAKIDPTDSKPEDMDK-PEHIPDPDAKPEDWDE 238
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 197 ysdwllppkkikdpsakpkedwdeqeylsdpdkkpgdyddipkeipdtsskpkedwdd 256
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 239 EMDCGEWEPVTONPEYKGEKPRQIDNDYKGTWHPHEDIDNPEYSPDPSTIAYDNFGLG 298
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 257 eedgewcaptlpnpeymgwkpkqkpnkpgkweapllidnpdfkddpelyvfpklkyvg 316
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 299 LDLMQVSGTIFDNFLITNDEAYAEFEFGNETGWVTKAAEKQMDKQDEEQRLKEEEDKK 358
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 317 lelqvksqslfndvlicddpdyakklatetwglkdaeka--afdaekkeeeeskd 373
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
QY 359 RKEEEAEADK-EDDE---DKDEDEDEDEKDEDEE---DVPGQAKDEL 400
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 374 apaesdaedepedgdgdsdeskaetksedsetsektdatahd 421
  |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|

RESULT 12
AAG24607
ID AAG24607 standard; Protein: 424 AA.
XX AC AAG24607;
XX AC AAG24607;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28349.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
99US-0132485.
99US-0132486.
99US-0132487.
99US-0132863.
99US-0134256.
99US-0134218.
99US-0134219.
99US-0134221.
99US-0134370.
99US-0134768.
99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138540.
99US-0138847.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139452.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
99US-0144334.
99US-0144335.
99US-0144352.
99US-0144632.
99US-0144884.
99US-0144814.
99US-0145086.
99US-0145088.
99US-0145085.

```

PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159299.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 52.8%; Score 1164; DB 21; Length 424;
 Best Local Similarity 54.2%; Pred. NO. 2.5e-90;
 Matches 221; Conservative 57; Mismatches 114; Indels 16; Gaps 8;

QY 3 AVYFEQFLDGDGWTSRWIESKHKSD---FGKFLSSGKFYGDDEKDKGLQTSODARFYA 59
 Db 23 avifeerf--ddgwenrwwksekddntagewkhtagnwsdg-andkgitsetsdyrfya 79
 QY 60 LSASFEPFSNKGQTLVVOFTVKHEQNIDCGGYVLFPSNLQDTDMHGDSSEYNIIMFGPDI 119
 Db 80 isaeifefsnkdktlvfqfsvkheqkldcggymkllsgdvqkkgfgdtpysimfgpdai 139
 QY 120 CGPGTKKHVIFNYKGNVINKDIRCKDDETHLYTLIVRPDNYEVKIDNSQVESGSL 179
 Db 140 cgystkkvhaityneanhlkkdvpcetdqthvytilrpdatsyllidnvektgsal 199
 QY 180 EDDWDFLPKIKIKDPDASKPEDWDERAKIDDPDTSKPEDMDK-PHIPPDPDAKKPEDWDE 238
 Db 200 ysdwllppkikidsakpedwdeqyisdpedkpgdyddipkeipdtcdskkpedwd 259
 QY 239 EMDGEWEPEVIONPEYKGEWKPRQIDNPDKGTWHPIDNPPEYSPDPSPYIAYDNFVGLG 298
 Db 260 eedgewtaptipnpeymgewkpkiknpnykqweaplidnpgdkddpelyvfpklyvg 319
 QY 299 LDLQWVKSTIFDNFLITNDEAYAEFEFNETWGVTKAEKQMKDKODEQRLKEEEDKK 358
 Db 320 lelqwksgslfdnvlidcddpdyakklatetwglkdaeka---afdeaeekneeeskd 376
 QY 359 RKEEEAEK-EDDE---DKDEDEDEDEKEDEEE--DVPQAKDEL 400
 Db 377 apaesdaedepedegdddseskaetksedsetsektahdel 424

RESULT 13
 AAG47932
 ID AAG47932 standard; Protein; 421 AA.
 XX
 AC AAG47932;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 60469.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999;	99US-0123180.
PR 09-MAR-1999;	99US-0123548.
PR 23-MAR-1999;	99US-0123788.
PR 25-MAR-1999;	99US-0126264.
PR 29-MAR-1999;	99US-0126785.
PR 01-APR-1999;	99US-0127462.
PR 06-APR-1999;	99US-0128234.
PR 08-APR-1999;	99US-0128714.
PR 16-APR-1999;	99US-0129845.
PR 19-APR-1999;	99US-0130077.
PR 21-APR-1999;	99US-0130449.
PR 23-APR-1999;	99US-0130510.
PR 28-APR-1999;	99US-0130891.
PR 28-APR-1999;	99US-0131449.
PR 30-APR-1999;	99US-0132048.
PR 30-APR-1999;	99US-0132407.
PR 04-MAY-1999;	99US-0132484.
PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 07-MAY-1999;	99US-0132487.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 24-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 21-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.

```

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      52.4%; Score 1157; DB 21; Length 421;
Best Local Similarity 53.9%; Pred. No. 9, 5e-90;
Matches 220; Conservative 57; Mismatches 115; Indels 16; Gaps 8;

QY 3  A Y F K E Q L D G G W T S R W T E S K H S D --- F G K F V L S S G K F Y G D E K D K G L Q T S Q D A R F Y A 59
Db 20  a v i e e r f - - d d g w e n r v w k s e w k d d n t a g e w k k t a g n w s g d - a n d k g i q t s e d y r f y a 76

QY 60  L S A S P E F S N K Q T L V Q F T V K H E Q N I D C G G Y V K L F P N S L D Q T D M H G D S E Y N I M F G P D I 119
Db 77  i s a e r p e f s n k d k t l v f g t s v k h e d k l d c g g y m k l l s g d v d q k f g g d t p y s i m f g p d i 136

QY 120  C G P G T K K V H I E N Y K G N V L I N K I R C K D D E F T H L Y T L I V R P D N T Y E V K I D N S Q V E S G S L 179
Db 137  c y s t k k v h a i l t y e a n h l i k k v p c e t d q l t h v t f i l r p d a t y s i l i d n v e k t g s l 196

QY 180  E D W D F L P P K K I D P A S K P E D W D E R A K I D D P T D S K P E D W D K - P E H I P D P A K K P E D W D E 238
Db 197  y s d w d l l p p k k i d p s a k k p e d w e q e y l s d p e d k k p d g y d i p e i p d t d s k k p e d w d d 256

QY 239  E M D G W E P P V I O N P E Y K E W K P R Q I D N D Y K T W H P E I D N P E Y S P D S I Y A Y D N F G V L G 298
Db 257  e e d g e w t a p t i p n e y m g e w k p k i k n p n y k g k w e a p l i d n p d f k d p e l y f p k l y v g 316

QY 299  L D L W O V K S G T I F D N F L I N D E A Y A E F G N E T W G V T K A A E K Q M K D Q D E O R L K B E E D K K 358
Db 317  l e l w q v g s g l d n v l i c d d p d y a k k l a d e t w g k l k d e a k a --- a f d e a e k n e e e s k d 373

QY 359  R K E E E A E D K - E D D E --- D K D E D E E D K E D E E E - - D V P Q A K D E L 400
Db 374  a p a e s d a e d e p d e g g d s d s e s k a e t k s v d s e e t s e k d a t a h d e l 421

```

RESULT 14
 AAG47931
 ID AAG47931 standard; Protein; 424 AA.
 XX
 AC AAG47931;
 XX

DT 18-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 60468.
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.

PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143342.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	26-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
<p>Query Match 52.4%; Score 1157; DB 21; Length 424; Best Local Similarity 53.9%; Pred. No. 9.6e-90; Matches 220; Conservative 57; Mismatches 115; Indels 16; Gaps 8;</p>					
QY	3	AVYFKEQFLDGDGWTSRWIESKHKSD---FCKFVLSGKGYGDEEKDKGLQTSODARFYA	59		
Db	23	avifeerf--ddgwenrwwksewkkddntagewkhtagnwsgd-andkgiqtsedyrfya	79		
QY	60	LSASFEPFSNKGQTLVVQFVVKHEONIDCGGYVVKLPNSLDQTDMDHGDSEYNIMFGPDI	119		
Db	80	isaeifefsnkdtlvfqsvkheqkdcggymkllsgdvdkkfggdtppsifmgpdi	139		
QY	120	CGPCTKKVHVIFNYKGNKLNKIDIRCKDDEFFHLYTLIVRPDNTYBVKIDNSOVESGSL	179		
Db	140	cgystkvhailtyneanhlkkadvpcetdqlthvyfllrpdatsyllidnvekqgsl	199		
QY	180	EDDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDMDK--PEHIPDPDAKPEDWDE	238		
Db	200	ysdwlllpkkikdpsakapedeqeyisdpedkkgdyddipkeipdtidskkpedwd	259		
QY	239	EMDCEWEPPIQNPYKGEWKPRDINDPDYKGTWHPIDNPYSPDPSIYAYDNFVLG	298		

Db 260 eedgewtaptipnpeymgwkqkqknpykgkweaplidnfpdkddpelyvfpklkyvg 319 PR 10-JUN-1999; 99US-0138540.
QY 299 LDIWQVKSCTIFDNFLITINDEAYAEFGNETWGVTKAAEKQMKRQKDEQRKKEEEDKK 358 PR 10-JUN-1999; 99US-0138847.
Db 320 lelwgqksgslfndvlicddpdyakkldetwglkdaeka---afdeaekknneeeskd 376 PR 16-JUN-1999; 99US-0139119.
QY 359 RKEEEAEUK-ENDE---DKDEDEDEEKEEEDDEE--DVPGQAKDEL 400 PR 16-JUN-1999; 99US-0139452.
Db 377 apaesdaedepegddgsdeskaetksvdssetsektahdel 424 PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

RESULT 15
AAG30997
ID AAG30997 standard; Protein: 441 AA.
XX
AC AAG30997;
DT
XX
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37155.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134258.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 14-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 19-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 21-AUG-1999; 99US-0149929.
PR 22-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 24-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 28-AUG-1999; 99US-0151080.
PR 29-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151438.
PR 31-AUG-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 02-SEP-1999; 99US-0153070.
PR 03-SEP-1999; 99US-0153758.
PR 04-SEP-1999; 99US-0154018.
PR 05-SEP-1999; 99US-0154039.
PR 06-SEP-1999; 99US-0154779.
PR 07-SEP-1999; 99US-0155139.
PR 08-SEP-1999; 99US-0155486.
PR 09-SEP-1999; 99US-0155659.
PR 10-SEP-1999; 99US-0156458.
PR 11-SEP-1999; 99US-0156596.
PR 12-SEP-1999; 99US-0157117.
PR 13-SEP-1999; 99US-0157753.
PR 14-SEP-1999; 99US-0157865.
PR 15-SEP-1999; 99US-0158029.
PR 16-SEP-1999; 99US-0158232.
PR 17-SEP-1999; 99US-0158369.
PR 18-SEP-1999; 99US-0159293.
PR 19-SEP-1999; 99US-0159294.
PR 20-SEP-1999; 99US-0159329.
PR 21-SEP-1999; 99US-0159330.
PR 22-SEP-1999; 99US-0159331.
PR 23-SEP-1999; 99US-0159637.
PR 24-SEP-1999; 99US-0159638.
PR 25-SEP-1999; 99US-0159584.
PR 26-SEP-1999; 99US-0160741.
PR 27-SEP-1999; 99US-0160767.
PR 28-SEP-1999; 99US-0160768.
PR 29-SEP-1999; 99US-0160770.
PR 30-SEP-1999; 99US-0160814.
PR 01-OCT-1999; 99US-0160815.
PR 02-OCT-1999; 99US-0160980.
PR 03-OCT-1999; 99US-0160981.
PR 04-OCT-1999; 99US-0160989.
PR 05-OCT-1999; 99US-0161404.
PR 06-OCT-1999; 99US-0161405.
PR 07-OCT-1999; 99US-0161406.
PR 08-OCT-1999; 99US-0161359.
PR 09-OCT-1999; 99US-0161360.
PR 10-OCT-1999; 99US-0161361.
PR 11-OCT-1999; 99US-0161920.
PR 12-OCT-1999; 99US-0161992.
PR 13-OCT-1999; 99US-0161993.
PR 14-OCT-1999; 99US-0162142.
PR 15-OCT-1999; 99US-0162142.
PR 16-OCT-1999; 99US-0162142.
PR 17-OCT-1999; 99US-0162142.
PR 18-OCT-1999; 99US-0162142.
PR 19-OCT-1999; 99US-0162142.
PR 20-OCT-1999; 99US-0162142.
PR 21-OCT-1999; 99US-0162142.
PR 22-OCT-1999; 99US-0162142.
PR 23-OCT-1999; 99US-0162142.
PR 24-OCT-1999; 99US-0162142.
PR 25-OCT-1999; 99US-0162142.
PR 26-OCT-1999; 99US-0162142.
PR 27-OCT-1999; 99US-0162142.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 51.8; Score 1142.5; DB 21; Length 441;
Best Local Similarity 53.3; Pred. No. 1.7e-88;
Matches 216; Conservative 57; Mismatches 109; Indels 23; Gaps 8;
QY 3 AVYFKEQFLDGDGWTNWSKHKSD--FGKFVLSGKFYGDDEKDKGLQTSQDARFYA 59

Db 20 avifeerf--ddgwenrwkxsewkkddntagewkhtegnwsgd-andkgigtseodyrfa 76
QY 60 LSASFEPFSNKGOTLVVQFTVKHEQNIDCGGYYVKLFPSNSLDQDTHMGDSSEYNIMFGPDI 119
Db 77 isaefpefsnkdktlvfgfsvkheqkldcggygmklsgdvdkkfggdtptysimfgpdi 136
QY 120 CGPOTKKVHVIFNYKGNVLINKDIRCKDDFTHTLYLVPRDNTYEVKIDNSQVESGSL 179
Db 137 cgystkvhalttyneanhlkkdvpcetdqlthvyflrpdatsilldnveqqtgsi 196
QY 180 EDDWDFLPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDK-PEHIFDPDAKPEDWDE 238
Db 197 ysdwllppkkikpsakpdedeqeyisdpedkpgdyddipkeipdtcdskpedwdd 256
QY 239 EMDGEWEPPIQNPPEYKGEWKPRQIDNPDKGTWIHPEIDNPESPDPSIYAYDNFVGLG 298
Db 257 eedgewtaptipnpeymgewkpkqikpnnykgkweaplidnpdfkddpelyvfpklkyvg 316
QY 299 LDLQWVKSCTIFDNFLITNDEAYAEFGNETWG-----VTKAAEKQMKQKQDEBQR 349
Db 317 lelwgksgslfdnvllicddpdyakkladetwglkdnlqsnitlqaeka---afdeaeak 373
QY 350 LKREEDKKRKEEREAEDK-EDDE--DKDEDEDEDEKDEDEE 390
Db 374 kneeeskdapaesdaedepdegdddseskaetksvdssee 418

Search completed: October 21, 2001, 03:05:46
Job time: 341 sec

THIS PAGE BLANK (USPTO)
